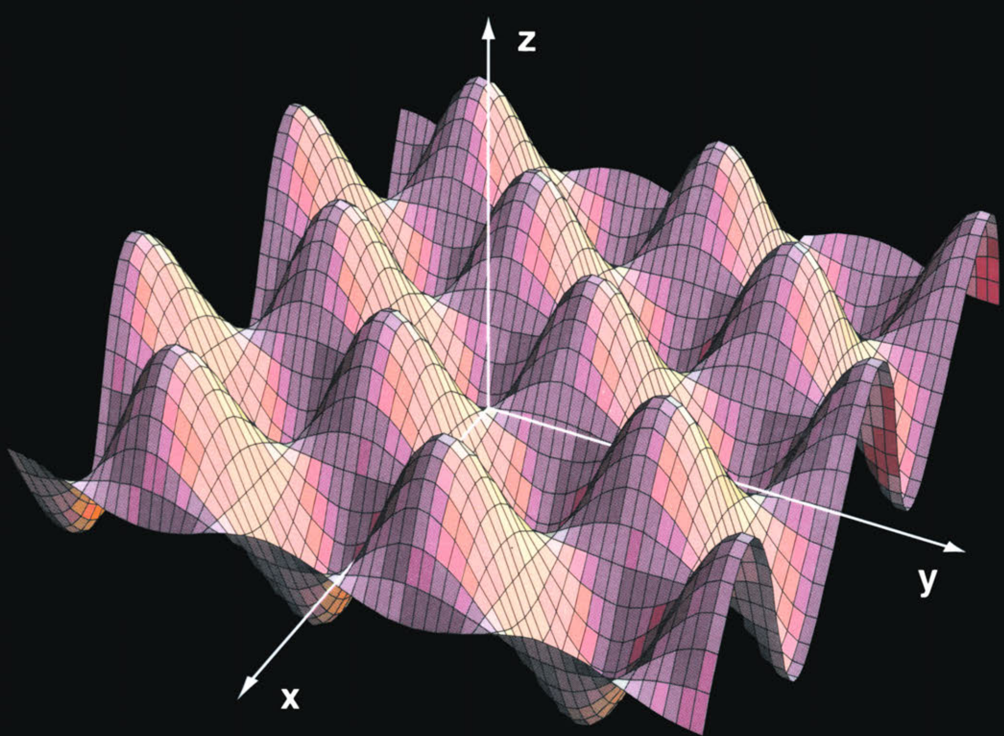
MACMILLAN COLLEGE WORK OUT SERIES

Advanced Calculus



Phil Dyke

MACMILLAN COLLEGE WORK OUT SERIES

Advanced Calculus

Titles in this Series

Advanced Calculus
Dynamics
Electric Circuits
Electronics
Elements of Banking
Engineering Materials
Fluid Mechanics
Heat and Thermodynamics
Mathematical Modelling Skills
Mechanics
Molecular Genetics
Numerical Analysis
Operational Research
Organic Chemistry
Physical Chemistry
Structural Mechanics

MACMILLAN COLLEGE WORK OUT SERIES

Advanced Calculus

Phil Dyke
Professor of Applied Mathematics
University of Plymouth, UK

First published 1998 by
MACMILLAN PRESS LTD
Houndmills, Basingstoke, Hampshire RG21 6XS
and London
Companies and representatives
throughout the world

ISBN 978-0-333-66092-8 ISBN 978-1-349-14076-3 (eBook)
DOI 10.1007/978-1-349-14076-3

A catalogue record for this book is available
from the British Library.

This book is printed on paper suitable for recycling and made from fully managed and sustained forest sources.

10 9 8 7 6 5 4 3 2 1
07 06 05 04 03 02 01 00 99 98

To my wife Heather

Contents

Preface

This Work Out is intended as a student guide to the applications of differential and integral calculus to vectors. Such topics usually form part of the later years of an engineering or (physical) applied science degree, or the second or later stages of the first year of a mathematics degree. Students in these categories will, I hope, benefit most from this book. For completeness, a chapter on vector algebra is also included even though this is not strictly calculus.

In the past few years there have been many changes in what is taught to the 17 to 19 age group who study mathematics. Most of these changes have been gradual; evolutionary rather than revolutionary. However, one trend that is universally acknowledged by those who teach is the diminution in the amount of time devoted to those parts of mathematics that depend upon the student being able to gain ability in algebraic manipulation. There have been many reasons given for this, but the major reason must be the change in both *how* mathematics is taught and *what* mathematics is taught to young children. The 13 to 16 year old in particular is no longer exposed to long sessions of drill examples in, for example, multiplying out brackets, factorising, simplifying or solving simultaneous and quadratic equations. There is so much more interesting mathematics to do: relating mathematics to the real world, modelling, introducing statistics and decision mathematics. The presence of sophisticated graphics calculators and algebraic manipulation packages also lends weight to those who are looking for an ally in their arguments to dispense with algebraic manipulation skills. While the existence of automatic ways to do algebra (and calculus too) is welcomed by the mathematics community, it is also recognised that there are dangers in making these tools available to the mathematically unsophisticated. Students who have always had access to automatic ways of doing algebra and calculus may lack the ability to carry out algebraic manipulation to the level required to do even quite simple mathematical problems. In particular, even the better students who possess a certain amount of mathematical talent can find the examples and exercises in this book trickier than they should because they lack the experience, and the speed that comes with that experience to undertake the necessary algebra to solve the problem. If a student gets stuck, or makes an error in the algebra, then the ability to 'see through to the end' of a problem can all but disappear. A pianist does not rely on a machine for the scales, he or she needs to know them before becoming competent. With competence comes confidence which is mandatory for a pianist and not far short of mandatory for a student of mathematics. In my experience, students get a great deal of satisfaction in carrying out manipulation accurately and efficiently, and being good at this side of mathematics is highly correlated to being good at other aspects such as having insight, knowing when to stop one approach and to try an alternative etc. Therefore in writing this book there is less emphasis on manipulation than in equivalent books written twenty or thirty years ago, and where it is unavoidable the student has the choice of doing the algebra by hand or using artificial aids. From the above, however, you will be able to deduce that the author prefers students to have at least some skill in algebraic manipulation and as you follow through examples you are encouraged to follow the mathematics carefully.

In this book, a knowledge of one-dimensional calculus has been assumed, although the first chapter provides a brief run through of the basic notions of limits, continuity and differentiability. It is *not* a rigorous approach, but more a recapitulation of some of the concepts and techniques required for the rest of the book. If any of the material in

Chapter 1 is too hard, then you need to study a first course on calculus before going any further. Much of what is in Chapter 1 is, however 'A'-level mathematics. In Chapter 2, the notion of a partial derivative is introduced, again from the point of view of a user of mathematics rather than someone who is overly concerned with proving all results rigorously. The results established in this chapter will be very useful for the rest of the book. Chapter 3 is devoted to finding extremes (stationary values, or 'maxima and minima') in two dimensions. Chapter 4 is a brief foray into the study of optimisation. The emphasis is on applying calculus as against the numerical aspects which can be found in texts on numerical methods. Vectors are introduced in Chapter 5. It has been assumed that students have no previous knowledge of vectors, although most will have at least a passing acquaintance with them which will be useful. Chapters 6 and 7 contain the application of the differential calculus to vectors and will be new or nearly new material for most students. The main application areas for vectors that change are in geometry, where there is a sub-branch called differential geometry devoted to describing curves and surfaces using vectors, and mechanics. Not very many mechanics examples feature here since there is a *Work Out Mechanics* devoted to the subject. Chapter 8 is a short introduction to line integrals which have applications in physics, especially electromagnetism (and mechanics). Chapter 9 is devoted to multiple integration which is the application of integration to areas and volumes. Applications here include finding areas, volumes, centres of mass and moments of inertia. Finally Chapters 10 and 11 deal with surface integrals and relations between line, surface and volume integrals. Again the applications here are in areas such as electromagnetism, finding areas of curved surfaces and mechanics.

In writing this, the author has been influenced by over ten years of teaching this subject to undergraduate mathematicians at the University of Plymouth and before that teaching engineering students. Most of the examples contained here have been tried out on students and have passed some sort of road test. It is a pleasure to acknowledge all of them for the feedback they have given over the years. I also thank all those who have provided examples. Finally a warm thanks to my family who did not have as much of my time as I would have liked through the production stages of this text.

PHIL DYKE
University of Plymouth
September 1996

Topic Guide

Limits
Continuity
Differentiability
Rules of Differentiation
L'Hôpital's Rule
Roots of Equations

1 Revision of One-Dimensional Calculus

1.1 Fact Sheet

In order to understand what a *derivative* is, the concepts of *function, limit* and *continuity* must first be understood. It will be assumed that the reader is at least acquainted with these notions from previous mathematical experience, however a quick run through of the concepts will be given here as a refresher.

A *function* is a rule that assigns precisely one value $f(x)$ to each element x, where x belongs to some set. In more everyday language, if x is a number, and f is a function then $f(x)$ is also a number; it is that number obtained by the insertion of the value x into the rule defined by f.

The concept of a *limit* is less straightforward. In the context of the calculus, attention is focused on what is meant by the expression $\lim_{x \to x_0} f(x) = L$ (alternatively written: $f(x) \to L$ as $x \to x_0$). If $f(x)$ is well defined at the point x_0 then the value of the limit is simply $L = f(x_0)$. The trouble comes when this is not the case, and the function $f(x)$ has no straightforward interpretation at $x = x_0$. In the calculus, the limit usually has the indeterminate form $\frac{0}{0}$ and various methods have to be used to calculate this (see Example 1.1). In this book, there will be no space devoted to the convergence of sequences or series, or the niceties of numerical analysis, all of which involve more subtle consideration of the concept of a limit.

Continuity is a concept that is central to the calculus, but once more there are many avenues and byways that cannot be explored here. These more analytical aspects are found in texts on Pure Mathematics specialising in analysis. Rigorous proofs of various results arising from the definition of continuity will not be found here. Instead consider the standard two-dimensional Cartesian set of axes x, y and a function $y = f(x)$. If the function f is continuous, then it has no breaks; a spider can crawl from one end to the other along its graph without having to leap or crawl infinitely fast! Slightly more rigorously, if a function f is defined in an interval $a \leq x \leq b$ where a and b are real numbers, and for every x_0 in this interval the $\lim_{x \to x_0} f(x)$ exists, is equal to $f(x_0)$ and is unique, then the function f is *continuous* at the point x_0, and since this can be any point in the interval $a \leq x \leq b$, we say that the function f is continuous in the interval $a \leq x \leq b$. The possibility of non-uniqueness arises because there are two directions from which x_0 can be approached, the left and the right. For a continuous function these two limits must be the same.

Having defined what is meant by a function, a limit and continuity it is now possible to define the *derivative*. If the $\lim_{x \to x_0} \left(\dfrac{f(x) - f(x_0)}{x - x_0} \right)$ exists and is unique, then the function f is said to be *differentiable* at x_0 or to possess a *derivative* at x_0. This quantity, called the derivative of f at $x = x_0$, is written as $\left.\dfrac{\mathrm{d}f}{\mathrm{d}x}\right|_{x=x_0}$ or more neatly as $f'(x_0)$ where the dash denotes the derivative of the function, and the symbol within the parentheses, x_0, the value of x where the derivative is to be evaluated. Higher derivatives are denoted by the number of dashes or perhaps a superfix in parentheses, so $f'''(x_0)$ denotes the third derivative of the function f at the point $x = x_0$, and $f^{(6)}(x_0)$ denotes the sixth derivative of f at the same location (etc.). To most readers, these definitions will come as no surprise. Just occasionally,

the letter f for the function and the notation $f(x)$ will be used interchangeably. This risks incurring the wrath of pure mathematicians everywhere! If any of this seems contrived and unfamiliar, then this book is probably too advanced and one of the plethora of rather thick books that usually rejoice under the name of *Calculus* is recommended as pre-reading. All that will be said here is that there are many practical applications of the differential calculus, those covered in this résumé are finding extrema, rates of change, gradients of tangents and numerical approximations.

The other subject of this revision chapter is *integration.* Integration is the inverse operation to differentiation. In other words, if a function $f(x)$ has a derivative $f'(x)$ (using the dash notation), then the function $f'(x)$ in turn has as its integral $f(x)$. Thus differentiation and integration have the same relation to each other as add and subtract or multiply and divide. (It is perhaps worth adding here that this inverse relationship is often obscured by the way many texts introduce integration in terms of finding the areas under graphs of functions. It is not easy to see that this has much to do with rates of change or slopes of tangents!) Once again, texts on the calculus are recommended to those for whom integration is a strange concept. In order to perform integration manually, specialist techniques such as substitution, integration by parts, the use of partial fractions or the application of reduction formulae are often required. These days, this is akin to manual long division, in that there is commercially available software for such tasks and they can be performed with ease and accuracy (computer algebra software). Only integration by parts is worth revising here, for use in future applications, and other techniques will not feature. Other applications of integration are the calculation of areas and volumes of revolution, centres of mass, and the solution of simple differential equations. Examples of some of these will be used as an aid to revision.

1.2 Worked Examples

Example 1.1 Find the value of each of the following limits (if it exists):

(a) $\lim_{x\to 2} (x^2 + 3x + 5)$,

(b) $\lim_{x\to 3} \left(\dfrac{x^3 - 2x^2 - x - 6}{x - 3}\right)$,

(c) $\lim_{x\to 0} \left(\dfrac{1 - \cos x}{x^2}\right)$,

(d) $\lim_{x\to 0} \left(\dfrac{\sqrt{x + 1} - 1}{x}\right)$,

(e) $\lim_{x\to 0} \dfrac{|x|}{x}$.

Solution This is the first example of the first chapter, and already it is time to part company with most texts on Pure Mathematics! In most cases, finding a limit is simply a case of inserting the value x_0 into the given formula for the function $f(x)$ to give $f(x_0)$, this being the unique value of $\lim_{x\to x_0} f(x)$. Pure mathematicians would take a great deal of trouble to establish that such limits exist and are unique, and to do so would make extensive use of a branch of mathematics called Real Analysis. This topic is usually a hard one for students, even mathematics students, but that is not a good reason for not pursuing it. A better reason, and the one used here, is that it is too much of a distraction at this stage. There *will* be some analysis in later chapters when genuinely new topics are introduced. So here, a heuristic (that is 'seat of pants') approach will be used to help the revision process. There are five limits to find here, which we shall tackle one by one.

(a) $\lim_{x\to 2} (x^2 + 3x + 5)$. In order to evaluate this limit, it really is merely a question of inserting the value $x = 2$ into the quadratic expression $x^2 + 3x + 5$ to obtain $2^2 + 3.2 + 5 = 15$. The value of the limit is thus 15.

(b) $\lim_{x\to 3}\left(\frac{x^3 - 2x^2 - x - 6}{x - 3}\right)$. When the value $x = 3$ is inserted into this expression, it takes the indeterminate form $\frac{0}{0}$. Whenever zero occurs in the denominator of an expression, we have what is called a *singularity*. Fortunately, all of the singularities encountered here are what is termed *removable*, that is they do not represent values of x for which the function $f(x)$ is infinite since there is always a compensating zero in the numerator at the crucial value. More about singularities can be found in textbooks on Complex Analysis. In the circumstances found in applications relevant to introducing the calculus, a method for evaluation needs to be used. Actually, indeterminate forms such as $\frac{0}{0}$, $\infty - \infty$, 1^∞ etc. can theoretically assume any value or even have no unique value. So this is another interface with Pure Mathematics which is circumvented here by actually evaluating the indeterminate forms, rather than proving general theorems about them. If this is done successfully, then the limit exists. If not, then it is usually because there is more than one choice for its value, so it certainly does not exist. In practical examples, it is usually one thing or the other and there is no uncertainty. In this example, it is convenient to use L'Hôpital's Rule, but this pre-supposes a knowledge of the differential calculus! Instead therefore factorisation will be used. Since the numerator, a cubic, is zero when $x = 3$, it must have $(x - 3)$ as a factor. By long division or other means of your choice, computer algebra perhaps, the following factorisation emerges: $x^3 - 2x^2 - x - 6 = (x - 3)(x^2 + x + 2)$. This means that the quotient inside the limit $\frac{x^3 - 2x^2 - x - 6}{x - 3} = x^2 + x + 2$. Thus $\lim_{x\to 3}\left(\frac{x^3 - 2x^2 - x - 6}{x - 3}\right) = \lim_{x\to 3}(x^2 + x + 2) = 14$.

(c) As in part (b), the expression $\frac{1 - \cos x}{x^2}$ takes the indeterminate form $\frac{0}{0}$ at $x = 0$. Once more, L'Hôpital's Rule could be applied but is not available. Another possibility is to use power series, but this also pre-supposes a knowledge of the calculus since the power series for $\cos x$ in ascending powers of x is simply Maclaurin's Series (Taylor's Series about $x = 0$). Instead, the trigonometric identity $1 - \cos x = 2\sin^2 \frac{1}{2}x$ is used together with the 'well known' limit $\lim_{x\to 0}\frac{\sin x}{x} = 1$. Evaluation of the limit thus proceeds as follows: $\lim_{x\to 0}\left(\frac{1 - \cos x}{x^2}\right) = \lim_{x\to 0}\left(\frac{2\sin^2 \frac{1}{2}x}{x^2}\right) = \frac{1}{2}\lim_{x\to 0}\left(\frac{\sin^2 \frac{1}{2}x}{(\frac{1}{2}x)^2}\right) = \frac{1}{2}\left(\lim_{u\to 0}\frac{\sin u}{u}\right)^2 = \frac{1}{2}$ where u has been written for $\frac{1}{2}x$.

(d) In order to find the value of $\lim_{x\to 0}\left(\frac{\sqrt{x + 1} - 1}{x}\right)$, the function in parentheses needs to be rationalised. This is done by using the relationship $\left(\sqrt{x + 1} - 1\right) = \left(\sqrt{x + 1} - 1\right)\frac{\left(\sqrt{x + 1} + 1\right)}{\left(\sqrt{x + 1} + 1\right)}$, which gives $\left(\sqrt{x + 1} - 1\right) = \frac{x + 1 - 1}{\sqrt{x + 1} + 1} = \frac{x}{\sqrt{x + 1} + 1}$. It is therefore apparent that $\lim_{x\to 0}\left(\frac{\sqrt{x + 1} - 1}{x}\right) = \lim_{x\to 0}\left(\frac{1}{\sqrt{x + 1} + 1}\right) = \frac{1}{2}$ since the last limit has no singularity at $x = 0$.

(e) If $x < 0$, $|x| = -x$, so $\frac{|x|}{x} = -1$ and since this expression is not dependent upon x, the value of the *left-hand* limit $\lim_{x\to 0_-}\frac{|x|}{x}$ must also be -1. For $x > 0$, $|x| = x$ so that this time, $\frac{|x|}{x} = 1$, and the *right-hand* limit $\lim_{x\to 0_+}\frac{|x|}{x}$ is 1. The notation used here is quite standard, $x \to 0_-$ means that x approaches zero from the left, and $x \to 0_+$ means that x approaches zero from the right. Since the values of these two limits are different (one is -1 and the other is $+1$), the conclusion is that the limit does not exist. It is worth making the point here that if a limit like this does not exist, then no calculus can be performed on the function. This is where Pure Mathematics comes into its own. For the user of mathematics (the engineer and applied scientist) has, in cases where the fundamentals break down, to rely on the pure mathematician to guide him or her through what can or cannot be done.

Example 1.2 The rigorous definition of a limit runs as follows: If for each $\varepsilon > 0$, $\exists$ a $\delta > 0$ such that $|f(x) - L| < \varepsilon$ whenever $|x - x_0| < \delta$ then we write $\lim_{x \to x_0} f(x) = L$. In this definition, the symbol $\exists$ means 'there exists'. Find a relationship between δ and ε given $f(x) = 5x - 7$, and $x_0 = 2$.

Solution If $x = 2$ then $f(2) = 3$, so we suspect that $L = 3$ in this case.

Now, $\quad |5x - 7 - 3| = |5x - 10| = 5|x - 2|$.

So if $\quad |x - 2| < \delta \quad$ then

$\quad |f(x) - L| = |5x - 7 - 3| = 5|x - 2| < 5\delta$.

So if $\quad \varepsilon = 5\delta$ we have $|f(x) - L| < \varepsilon$ whenever $|x - x_0| < \delta$.

Hence $\delta = \dfrac{\varepsilon}{5}$. Figure 1.1 displays this relationship.

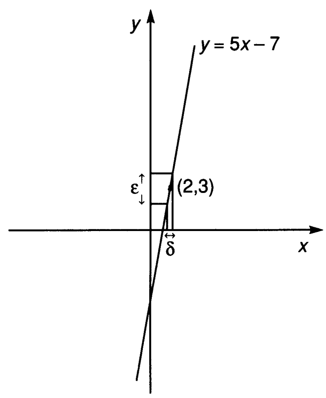


Figure 1.1 The relationship between ε and δ

Example 1.3 Show that $\lim_{x \to 0} \left(x \sin \dfrac{1}{x} \right) = 0$.

Solution There is a choice of method for solving this particular problem, but it is constructive to choose one that demonstrates a useful result sometimes known as the *pinching theorem*. This theorem states that, if $a(x) \leq f(x) \leq b(x)$ for some three functions $a(x)$, $b(x)$ and $f(x)$ suitably well defined in intervals that all contain the value $x = x_0$, and if $\lim_{x \to x_0} a(x) = L$ and $\lim_{x \to x_0} b(x) = L$, then it must be the case that $\lim_{x \to x_0} f(x) = L$ too. Figure 1.2(a) shows in pictorial form why this theorem holds, and why it is called the pinching theorem.

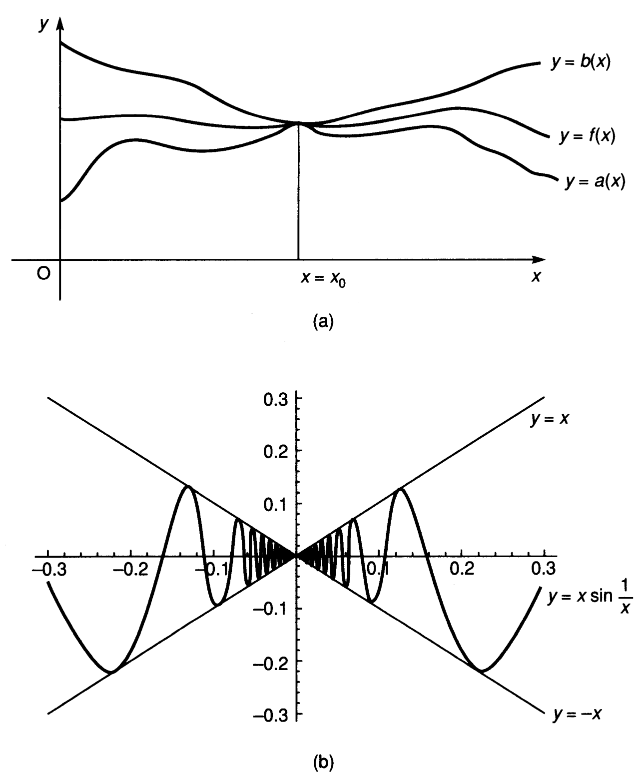


Figure 1.2 (a) The pinching theorem in general. (b) The pinching theorem for $f(x) = x \sin\left(\dfrac{1}{x}\right)$.

Let us now use it to solve the limit given in the question. The sine function always lies between -1 and $+1$ (for real variables), so $\left|\sin\frac{1}{x}\right| \leq 1$ for all x. This implies that $\left|x\sin\frac{1}{x}\right| = |x|\left|\sin\frac{1}{x}\right| \leq |x|$. Now, since $0 \leq \left|x\sin\frac{1}{x}\right| \leq |x|$, and $|x| \to 0$ as $x \to 0$, by the pinching theorem, $\lim_{x\to 0}\left|x\sin\frac{1}{x}\right| = 0$. Since $-\left|x\sin\frac{1}{x}\right| \leq x\sin\frac{1}{x} \leq \left|x\sin\frac{1}{x}\right|$ using the result just obtained together with the pinching theorem again, proves that $\lim_{x\to 0}\left(x\sin\frac{1}{x}\right) = 0$. Figure 1.2(b) shows this particular case of the theorem.

Example 1.4 £1000 is invested in an account that pays 7 per cent interest compounded n times each year; this means that there will be $1000\left(1 + \frac{0.07}{n}\right)^{10n}$ pounds in 10 years. Find how much money will be in the account after 10 years if:

(a) $n = 4$ (quarterly investment),
(b) $n = 12$ (monthly investment),
(c) $n = 365$ (daily investment).
(d) What is the theoretical maximum amount of money that could be in the account after 10 years?

Solution This example is an economic application of limits. The function $f(n) = 1000\left(1 + \frac{0.07}{n}\right)^{10n}$ is evaluated for parts (a), (b) and (c). This is straightforward, and the figures are:

(a) $f(4) = 1000\left(1 + \frac{0.07}{4}\right)^{10.4} = £2001.60$,

(b) $f(12) = 1000\left(1 + \frac{0.07}{12}\right)^{10.12} = £2009.60$,

(c) $f(365) = 1000\left(1 + \frac{0.07}{365}\right)^{10.365} = £2013.60$.

(d) The formula for calculating the amount of money in the account after 10 years is $f(n) = 1000\left(1 + \frac{0.07}{n}\right)^{10n}$ and the previous calculations lead us to suspect that this sequence increases with n. In fact $\left(1 + \frac{a}{n}\right)^{bn}$ can be expanded using the binomial theorem as follows: $\left(1 + \frac{a}{n}\right)^{bn} = 1 + bn \cdot \frac{a}{n} + \frac{bn(bn-1)}{2}\left(\frac{a}{n}\right)^2 + \cdots + \frac{bn(bn-1)(bn-2)\cdots(bn-k+1)}{k!}\left(\frac{a}{n}\right)^k + \cdots$

The general term is $\dfrac{b\left(b - \frac{1}{n}\right)\left(b - \frac{2}{n}\right)\cdots\left(b - \frac{(k-1)}{n}\right)}{k!}a^k$ which $\to \frac{(ba)^k}{k!}$ as $n \to \infty$. Thus $\left(1 + \frac{a}{n}\right)^{bn} \to 1 + ab + \cdots + \frac{(ab)^k}{k!} + \cdots = e^{ab}$ as $n \to \infty$. Hence as $n \to \infty$, $f(n) = 1000\left(1 + \frac{0.07}{n}\right)^{10n} \to 1000e^{0.07\times 10} = 2014$. Our suspicions are confirmed, and the sequence is indeed increasing giving £2014 as the maximum amount that could possibly accrue after 10 years (by *continuously* reinvesting, a practical impossibility well approximated by reinvesting daily, part (c)).

Example 1.5 The limit $\lim_{t\to t_0}\left(\frac{x(t) - x(t_0)}{t - t_0}\right)$ denotes the speed of a particle whose distance x is defined in terms of time t by the function $x(t)$. Find the speed of the particle at the time t_0 for the following three cases:

(a) $x = t^2 + 2$,
(b) $x = \sin t$,
(c) $x = t^n$, where n is a constant.

Solution (a) For this case, the limit $\lim_{t\to t_0}\left(\frac{x(t) - x(t_0)}{t - t_0}\right)$ is given by $\lim_{t\to t_0}\left(\frac{t^2 - t_0^2}{t - t_0}\right) = \lim_{t\to t_0}\left(\frac{(t - t_0)(t + t_0)}{t - t_0}\right) = \lim_{t\to t_0}(t + t_0) = 2t_0$.

(b) For this second case, the limit $\lim_{t\to t_0}\left(\frac{x(t) - x(t_0)}{t - t_0}\right)$ is now given by $\lim_{t\to t_0}\left(\frac{\sin t - \sin t_0}{t - t_0}\right)$. This is not quite so easy to simplify, and demands knowledge of the trigonometric formula usually written

$$\sin A - \sin B = 2\sin\left(\frac{A-B}{2}\right)\cos\left(\frac{A+B}{2}\right)$$

Applying this formula leads to the following evaluation of the limit:

$$\lim_{t\to t_0}\left(\frac{\sin t - \sin t_0}{t - t_0}\right) = \lim_{t\to t_0}\left(\frac{2\sin\frac{1}{2}(t-t_0)\cos\frac{1}{2}(t+t_0)}{t-t_0}\right)$$

$$= \lim_{u\to 0}\frac{\sin u}{u}\lim_{t\to t_0}\cos\tfrac{1}{2}(t+t_0) = \cos t_0$$

In this last step, u has been written instead of $\frac{1}{2}(t - t_0)$ and the following property of limits has also been utilised: $\lim_{x\to x_0}(f(x)\,g(x)) = \lim_{x\to x_0}f(x)\lim_{x\to x_0}g(x)$ (provided both limits exist and are unique). The fact that $\frac{\sin u}{u} \to 1$ as $u \to 0$ has also been used again.

There are many standard rules obeyed by limits, most of them are obvious, but some of them less so. They have varying degrees of usefulness, and those necessary will be introduced when required. It goes without saying that those calculus books already mentioned will usually contain more about limit theorems.

(c) In order to evaluate $\lim_{t\to t_0}\left(\frac{t^n - t_0^n}{t - t_0}\right)$, we first write $t - t_0 = h$ so that the limit becomes $\lim_{h\to 0}\left(\frac{(t_0+h)^n - t_0^n}{h}\right)$. It is now possible to use the binomial theorem to find the value of this limit. Now, $(t_0 + h)^n = t_0^n\left(1 + \frac{h}{t_0}\right)^n = t_0^n\left(1 + \frac{nh}{t_0} + \frac{n(n-1)}{2}\left(\frac{h}{t_0}\right)^2 + \cdots\right)$ so that $(t_0 + h)^n - t_0^n = nht_0^{n-1} + O(h^2)$ where the symbol $O(h^2)$ denotes terms containing h^2 and higher powers. Hence, $\lim_{h\to 0}\left(\frac{(t_0+h)^n - t_0^n}{h}\right)$ becomes nt_0^{n-1}. This is the required result.

Of course, what has been established here are three **derived functions** or **derivatives**. The last two results, which in more normal notation are $\frac{d}{dx}(\sin x) = \cos x$ and $\frac{d}{dx}(x^n) = nx^{n-1}$ are particularly useful. No further results of this kind will be derived, as they are all on commonly available formulae sheets and computer algebra software.

Example 1.6 Establish the following rules of differentiation:

(a) $\frac{d}{dx}(cf(x)) = c\frac{df}{dx}$,

(b) $\frac{d}{dx}(f(g(x))) = \frac{df}{dg}\frac{dg}{dx}$ (the function of a function rule),

(c) $\frac{d}{dx}(u(x)v(x)) = u(x)\frac{dv}{dx} + v(x)\frac{du}{dx}$ (the product rule),

(d) $\frac{d}{dx}\left(\frac{u(x)}{v(x)}\right) = \frac{v\frac{du}{dx} - u\frac{dv}{dx}}{v^2}$ (the quotient rule).

Solution All of these rules are derived from the limit definition of the derivative. It is most convenient to use the following form: $\frac{df}{dx} = \lim_{h\to 0}\left(\frac{f(x+h) - f(x)}{h}\right)$.

(a)
$$\frac{d}{dx}(cf(x)) = \lim_{h\to 0}\left(\frac{cf(x+h) - cf(x)}{h}\right)$$

$$= c\lim_{h\to 0}\left(\frac{f(x+h) - f(x)}{h}\right)$$

$$= c\frac{df}{dx} \text{ as required}$$

(b) $\dfrac{\mathrm{d}}{\mathrm{d}x}(f(g(x))) = \lim_{h\to 0}\left(\dfrac{f(g(x+h)) - f(g(x))}{h}\right) = \lim_{h\to 0}\left[\left(\dfrac{f(g(x+h)) - f(g(x))}{g(x+h) - g(x)}\right)\left(\dfrac{g(x+h) - g(x)}{h}\right)\right]$

$$= \lim_{k\to 0}\left(\frac{f(g+k) - f(g)}{k}\right)\lim_{h\to 0}\left(\frac{g(x+h) - g(x)}{h}\right)$$

$$= \frac{\mathrm{d}f}{\mathrm{d}g}\frac{\mathrm{d}g}{\mathrm{d}x}$$

Where in this example, $k = g(x + h) - g(x)$, and $k \to 0$ as $h \to 0$ since $g(x)$ is a continuous function. Thus $k \to 0$ can replace $h \to 0$ in the first limit.

(c) $\dfrac{\mathrm{d}}{\mathrm{d}x}(u(x)\,v(x)) = \lim_{h\to 0}\left(\dfrac{u(x+h)\,v(x+h) - u(x)\,v(x)}{h}\right)$

$$= \lim_{h\to 0}\left(\frac{u(x+h)\,v(x+h) - v(x+h)\,u(x) + v(x+h)\,u(x) - u(x)\,v(x)}{h}\right)$$

$$= \lim_{h\to 0}\left[\left(\frac{u(x+h) - u(x)}{h}\right)v(x+h)\right] + \lim_{h\to 0}\left[\left(\frac{v(x+h) - v(x)}{h}\right)u(x)\right]$$

$$= v(x)\frac{\mathrm{d}u}{\mathrm{d}x} + u(x)\frac{\mathrm{d}v}{\mathrm{d}x} \text{ as required}$$

(d) In order to establish the quotient rule, we combine results (b) and (c) together with the standard form $\dfrac{\mathrm{d}}{\mathrm{d}x}\left(\dfrac{1}{x}\right) = -\dfrac{1}{x^2}$.

Writing $w = \dfrac{1}{v}$, we have just proved $\dfrac{\mathrm{d}}{\mathrm{d}x}(uw) = w\dfrac{\mathrm{d}u}{\mathrm{d}x} + u\dfrac{\mathrm{d}w}{\mathrm{d}x}$. Also, $\dfrac{\mathrm{d}w}{\mathrm{d}x} = \dfrac{\mathrm{d}}{\mathrm{d}x}\left(\dfrac{1}{v}\right) = \dfrac{\mathrm{d}v}{\mathrm{d}x}\dfrac{\mathrm{d}}{\mathrm{d}v}\left(\dfrac{1}{v}\right) = -\dfrac{1}{v^2}\dfrac{\mathrm{d}v}{\mathrm{d}x}$ where we have used (b) in the second equality. Combining these two leads to $\dfrac{\mathrm{d}}{\mathrm{d}x}\left(\dfrac{u}{v}\right) = \dfrac{1}{v}\dfrac{\mathrm{d}u}{\mathrm{d}x} - \dfrac{u}{v^2}\dfrac{\mathrm{d}v}{\mathrm{d}x}$, that is $\dfrac{\mathrm{d}}{\mathrm{d}x}\left(\dfrac{u}{v}\right) = \dfrac{v\dfrac{\mathrm{d}u}{\mathrm{d}x} - u\dfrac{\mathrm{d}v}{\mathrm{d}x}}{v^2}$, the required result.

Example 1.7 Establish whether the function $f(x)$ defined by

$f(x) = \begin{cases} x^2 & \text{if } x \le 0 \\ -x^2 & \text{if } x > 0 \end{cases}$ has a derivative at $x = 0$.

Solution In order to solve this particular problem, the limit of the expression $\dfrac{f(x) - f(0)}{x - 0}$ as x tends to 0 needs to be investigated. We therefore proceed as follows: Approaching 0 from the left, $\lim_{x\to 0_-}\left(\dfrac{f(x) - f(0)}{x - 0}\right) = \lim_{x\to 0_-}\dfrac{x^2}{x} = \lim_{x\to 0_-} x = 0$. Approaching 0 from the right, $\lim_{x\to 0_+}\left(\dfrac{f(x) - f(0)}{x - 0}\right) = \lim_{x\to 0_+}\dfrac{-x^2}{x} = \lim_{x\to 0_+}(-x) = 0$. These limits are the same, hence the limit is unique and the derivative is indeed well defined at $x = 0$ with the value 0. In reality, this means that, despite its piecemeal definition, the function $f(x)$ at $x = 0$ is smooth (has no kinks) and possesses a unique tangent. See Figure 1.3.

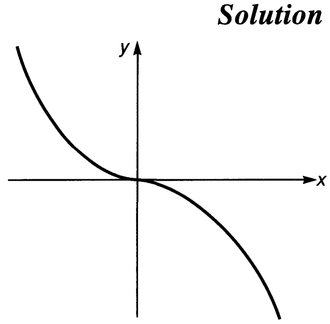


Figure 1.3 The function $f(x) = \begin{cases} x^2 & x \le 0 \\ -x^2 & x > 0 \end{cases}$

Example 1.8 Demonstrate with the use of diagrams the truth of the following two important results:

(a) Rolle's Theorem: If $f(x)$, $(a \le x \le b)$ is a differentiable function and $f(a) = f(b)$, then there is at least one point $x = c$ in (a, b) such that $f'(c) = 0$.

(b) The First Mean Value Theorem (of the Differential Calculus): If the function $f(x)$ is continuous on the domain $[a, b]$ and differentiable on (a, b) then there exists at least one point $x = c$, with $a < c < b$, such that $\dfrac{f(b) - f(a)}{b - a} = f'(c)$.

Solution Both of these results play a central role in the development of the Differential Calculus, but their rigorous proofs depend too much upon analysis which is out of keeping with the applied tone of this text. Instead, their truth is demonstrated in such a way that, hopefully, there is no doubt about their validity. Rolle's Theorem is demonstrated clearly through Figure 1.4. This shows that if a function is smooth (differentiable) and its end values are the same, then there must be a point somewhere in the interval where the slope of the tangent (its derivative) is zero.

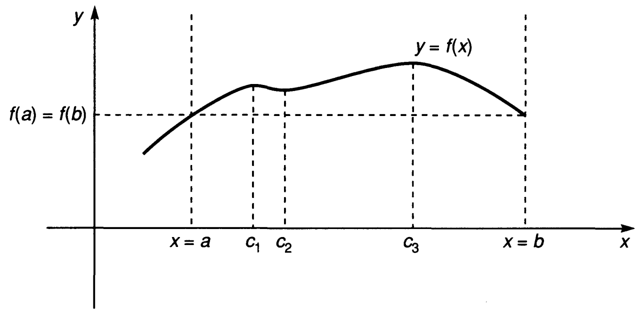


Figure 1.4 It is not possible for $f(a) = f(b)$ without there being a value c such that $f'(c) = 0$, $a < c < b$ provided $f(x)$ is differentiable. In this figure, there are *three* values of c (labelled c_1, c_2 and c_3).

The Mean Value Theorem is a slight generalisation of this, in that (see Figure 1.5) it implies that if a function is smooth between two values a and b, then there must be a value of the slope of the tangent somewhere between these two values which is the same as the slope of the chord joining the end points a and b.

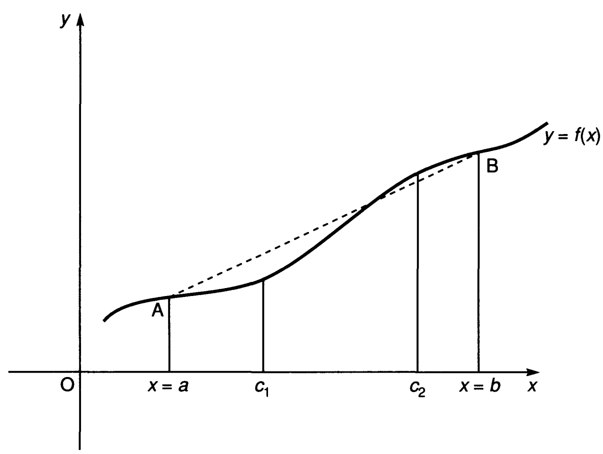


Figure 1.5 The Mean Value Theorem. The line AB has a slope that is equalled by the slope of the tangent to a point $a < c < b$ at least once provided $f(x)$ is differentiable on $[a, b]$. There are two such points in this figure, labelled c_1 and c_2.

Note the notation of the question which uses square brackets to denote the interval that includes the end points, and parentheses to denote the interval that excludes them. This is entirely standard notation.

Example 1.9 Show that $f(b) = f(a) + f'(a)(b - a) + f''(c)\dfrac{(b-a)^2}{2!}$ where $f(x)$ is a suitably well behaved function (that is, it possesses a second derivative in the interval (a, b)) and c is a point within this interval.

Solution This solution is rather artificial, in that it depends on being able to dream up the function: $g(x) = f(b) - f(x) - f'(x)(b - x) - (b - x)^2 A$, where A is a constant chosen carefully so that $g(a) = 0$. It is already true that $g(b) = 0$, so once $g(a) = 0$, the conditions of Rolle's Theorem apply, and c is the point in the interval (a, b) where the derivative of $g(x)$ vanishes, that is $g'(c) = 0$. Doing the algebra, demanding that $g(a) = 0$ gives $0 = f(b) - f(a) - f'(a)(b - a) - (b - a)^2 A$ so that $A = \dfrac{f(b) - f(a) - f'(a)(b-a)}{(b-a)^2}$. Now, $g'(c) = 0$ gives, on differentiating the expression for $g(x)$, $g'(c) = 0 = -f'(c) - f''(c)(b - c) + f'(c) + 2(b - c)A$ or $A = \frac{1}{2}f''(c)$. Finally, substituting for this value of A gives $\frac{1}{2}f''(b) = \dfrac{f(b) - f(a) - f'(a)(b-a)}{(b-a)^2}$ which, when rearranged as a formula

for $f(b)$ gives the desired result $f(b) = f(a) + f'(a)(b - a) + f''(c)\dfrac{(b-a)^2}{2!}$. This proof may not seem particularly illuminating, but it does lead, by generalisation, to one of the most important theorems of the calculus, Taylor's Theorem:

$$f(b) = f(a) + f'(a)(b - a) + f''(a)\frac{(b-a)^2}{2!} + f'''(a)\frac{(b-a)^3}{3!} + \cdots + f^{(n)}(a)\frac{(b-a)^n}{n!} + R_n$$

where the dashes against $f(x)$ (and the (n)) denote the order of the derivative, and the final term R_n denotes the remainder term, usually very small and of little practical use except in Numerical Analysis, and containing the $(n + 1)$th derivative of $f(x)$ evaluated at some point in the interval (a, b). Taylor's Theorem can be thought of as an nth Mean Value Theorem, giving a very accurate approximation to $f(b)$ in terms of f and all its derivatives at $x = a$. The nearer a is to b, the better the approximation and the fewer derivatives of f that need to be calculated (in general). Of course, the function $f(x)$ has to be appropriately well behaved in the interval $[a, b]$, that is continuous on the closed interval and many $(n + 1)$ times differentiable in the open interval (a, b).

Example 1.10 L'Hôpital's Rule runs as follows: Suppose 'lim' denotes one of the symbols $\lim\limits_{x\to x_0}$, $\lim\limits_{x\to x_0-}$, $\lim\limits_{x\to x_0+}$, $\lim\limits_{x\to\infty}$, $\lim\limits_{x\to-\infty}$ and the functions $f(x)$ and $g(x)$ are differentiable where defined, except possibly at $x = x_0$. If $\lim f(x) = \lim g(x) = 0$, then $\lim \dfrac{f(x)}{g(x)} = \lim \dfrac{f'(x)}{g'(x)}$, provided the limit on the right exists or is infinite. Prove this rule, and use it to determine the following limits:

(a) $\lim\limits_{x\to 1}\left(\dfrac{1-x}{e^x - e}\right)$,

(b) $\lim\limits_{x\to\pi/2}\left(\dfrac{1-\sin x}{\cos x}\right)$,

(c) $\lim\limits_{x\to 0}\left(\dfrac{\sin x - x}{x - \tan x}\right)$,

(d) $\lim\limits_{x\to 1}\left(\dfrac{x^{\frac{1}{2}} - x^{\frac{1}{4}}}{x-1}\right)$.

Solution In order to prove this rule (which incidentally seems to have been discovered, not by the Frenchman G. F. A. L'Hôpital (1661–1704) but by his more illustrious teacher Johann Bernoulli (1667–1748)) Taylor's Theorem will be used. First of all, suppose that the zeros of both $f(x)$ and $g(x)$ occur at $x = x_0$; we can then expand each function in terms of its Taylor's Series as follows: $\dfrac{f(x)}{g(x)} = \dfrac{f(x_0) + (x - x_0) f'(x_0) + \frac{1}{2}(x - x_0)^2 f''(x_0) + \cdots}{g(x_0) + (x - x_0) g'(x_0) + \frac{1}{2}(x - x_0)^2 g''(x_0) + \cdots}$. Note that the first terms in both numerator and denominator are zero, cancel a factor $(x - x_0)$ then let $x \to x_0$. Provided $f'(x_0)$ and $g'(x_0)$ are not both zero, we have, immediately, $\lim\limits_{x\to x_0}\dfrac{f(x)}{g(x)} = \lim\limits_{x\to x_0}\dfrac{f'(x)}{g'(x)}$. If $f'(x_0) = 0$, but $g'(x_0) \neq 0$, the limit is zero, and if $f'(x_0) \neq 0$, but $g'(x) = 0$, the limit is infinity. If both first derivatives are zero, then the first *two* terms of both numerator and denominator are zero, and the limit $\lim\limits_{x\to x_0}\dfrac{f(x)}{g(x)}$ is now equal to $\lim\limits_{x\to x_0}\dfrac{f''(x)}{g''(x)}$ after cancellation of the factor $\frac{1}{2}(x - x_0)^2$. If the zero at $x = x_0$ is of even higher order, more terms of the Taylor's Series of each function are required, but it is important to remember that *both* $f(x)$ and $g(x)$ (and derivatives) need to be zero before L Hôpital's Rule can be applied. If just one is zero, the cancellations of the factors containing $(x - x_0)$ does not happen and the rule is invalid. This is a particularly common source of error among engineers and applied scientists.

(a) The limit $\lim\limits_{x\to 1}\left(\dfrac{1-x}{e^x - e}\right)$, is of the form $\dfrac{0}{0}$, hence L'Hôpital's Rule is applicable. Differentiating top and bottom gives $\dfrac{-1}{e^x}$, and letting $x \to 1$ yields the limit $-1/e$.

(b) The limit $\lim\limits_{x\to\pi/2}\left(\dfrac{1-\sin x}{\cos x}\right)$ is also of the form $\dfrac{0}{0}$, so proceeding to differentiate top and bottom as before, the new quotient becomes $\dfrac{-\cos x}{-\sin x}$ which clearly tends to zero as $x \to \frac{1}{2}\pi$. The value of

the limit is therefore 0. This is one of those cases that need watching as the unwary may continue to differentiate, forgetting that 0 is a perfectly allowable limit.

(c) The limit $\lim_{x\to 0}\left(\dfrac{\sin x - x}{x - \tan x}\right)$ is also amenable to L'Hôpital's Rule, but the resulting quotient on differentiating top and bottom is $\dfrac{\cos x - 1}{1 - \sec^2 x}$ which is still $\dfrac{0}{0}$. We are therefore allowed to differentiate again, this time obtaining $\dfrac{-\sin x}{-2\sec x\ \tan x}$ which although it is yet again an indeterminate form, is in fact $\dfrac{\sin x}{2\dfrac{\sin x}{\cos^2 x}} = \frac{1}{2}\cos^2 x = \frac{1}{2}$ when $x = 0$. The value of the limit is thus $\frac{1}{2}$.

(d) The limit $\lim_{x\to 1}\left(\dfrac{x^{\frac{1}{2}} - x^{\frac{1}{4}}}{x - 1}\right)$ is the same as $\lim_{x\to 1}\left(\dfrac{\frac{1}{2}x^{-\frac{1}{2}} - \frac{1}{4}x^{-\frac{3}{4}}}{1}\right)$ (on differentiating top and bottom) which, straightforwardly, is $\frac{1}{4}$.

Example 1.11 Show that, if $x = x_n$ is an approximation to a root of the equation $f(x) = 0$, then a better approximation is given by

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

provided $f'(x_n) \neq 0$. Use this formula (called the Newton–Raphson Method) to approximate the cube root of ten to five decimal places.

Solution Taylor's Theorem, retaining only the first two terms, is

$$f(x + h) \approx f(x) + hf'(x)$$

Letting $x = x_n$ and $x + h = x_{n+1}$ leads to

$$f(x_{n+1}) = f(x_n) + hf'(x_n)$$

Let $f(x_{n+1}) = 0$, that is x_{n+1} is the exact root, therefore

$$h = -f(x_n)/f'(x_n)$$

so

$$x_{n+1} = x_n + h$$

gives the result known as the Newton–Raphson Formula:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

The ≈ has been replaced by = since approximation is understood. Also all terms on the right are known (since x_n is known) hence x_{n+1} can be found. This is an example of an *explicit* formula. If x_{n+1} also occurred on the right, the formula would be *implicit* and consequently less useful. Figure 1.6 shows how this formula gives successive approximations that (in general) home in on a root.

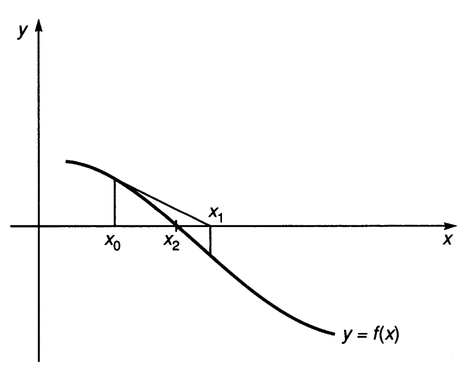


Figure 1.6 The Newton–Raphson Method. Here, only three iterations get very close to the root of $f(x) = 0$ (at x_2).

In this figure, $\tan\theta = f'(x_n)$, the derivative being the slope of the tangent. However, also from the diagram, $\tan\theta = \dfrac{f(x_n)}{x_n - x_{n+1}}$. Equating these two expressions for $\tan\theta$ re-derives the Newton–Raphson Formula. Looking at Figure 1.6, it can be seen that the Newton–Raphson Method for calculating roots of equations can break down if the gradient of the solution curve $y = f(x)$ varies wildly. In particular if there is a value of x for which $f'(x) = 0$ lying between x_n and the true zero. Now we are ready to see how the method works.

The cube root of ten is the positive root of the equation $x^3 - 10 = 0$. To find an approximation, let us tabulate a few values of $f(x) = x^3 - 10$ and find where it changes sign:

$f(0)$	$f(1)$	$f(2)$	$f(3)$
-10	-9	-2	17

The root is thus between $x = 2$ and $x = 3$. Hence, as a first approximation, choose $x_0 = 2$. Differentiating gives $f'(x) = 3x^2$ and this tells us that $f'(x)$ is not close to zero, hence we expect the method to work well. In general, a large value of $f'(x_n)$ gives rapid convergence of the method as is the case here ($f'(x_n) \geq 12$). The following table gives a systematic way of carrying out the Newton–Raphson process:

x	$f(x)$	$f'(x)$	$f(x)/f'(x)$	New x
2	-2	12	-0.166667	2.166667
2.166667	0.171297	14.08333	0.012163	2.154504
2.154504	0.000960	13.92566	0.000069	2.154435
2.154435	0	13.92477	0	2.154435

It is obvious that the last number in the 'New x' column is the required cube root of 10, that is 2.15444 to 5 decimal places. Note that the method naturally terminates to the number of decimal places required as the quotient $f(x)/f'(x)$ is zero in the last row.

Example 1.12 Find and classify the extreme values of the following functions:

(a) $f(x) = x^2 + \dfrac{1}{x}$,

(b) $f(x) = (x - 1)^2(x - 2)^2$,

(c) $f(x) = \dfrac{x^2}{1 + x^2}$,

(d) $f(x) = \sin^2 x - \sqrt{3}\cos x,\ 0 \leq x \leq \pi$.

Solution This kind of problem is solved by finding the first derivative and setting it equal to zero in order to determine any extreme values. There are three species of extreme; minimum, maximum and point of inflection, and these are distinguished as follows. A minimum value is a point that is surrounded by function values larger than itself, and is tested formally by examining the sign of the *second* derivative. If this second derivative, $\dfrac{d^2f}{dx^2}$ is positive at the extremum it is a minimum. A maximum value is surrounded by function values smaller than itself, and at a maximum, the second derivative is negative. A point of inflection is an extremum that is neither a maximum nor a minimum and at such a point, $\dfrac{d^2f}{dx^2} = 0$. However, if $\dfrac{d^2f}{dx^2} = 0$ then we do not know whether the extremum is a maximum, minimum or point of inflection. This is a classic case of a *necessary* but not *sufficient* condition. Let us solve the four problems.

(a) If $f(x) = x^2 + \dfrac{1}{x}$, then $f'(x) = 2x - \dfrac{1}{x^2}$, and this is zero at $x = 2^{-\frac{1}{3}}$. At this value $f''(x) = 2 + \dfrac{2}{x^3} = 6 > 0$, so $x = 2^{-\frac{1}{3}}$ is a minimum.

(b) If $f(x) = (x - 1)^2(x - 2)^2$, then $f'(x) = 2(x - 1)^2(x - 2) + 2(x - 1)(x - 2)^2$, and this is zero at the three values $x = 1, \frac{3}{2}, 2$. The second derivative is given by $f''(x) = 2(x - 1)^2 + 8(x - 1)(x - 2) + 2(x - 2)^2$, and this expression is positive at $x = 1$ and $x = 2$, but negative at $x = \frac{3}{2}$. Thus $x = 1$ and $x = 2$ are both minima, while $x = \frac{3}{2}$ is a maximum.

(c) If $f(x) = \dfrac{x^2}{1 + x^2}$, then $f'(x) = \dfrac{2x(1 - x^2) - 2x.x}{(1 + x^2)^2}$, which is only zero if $x = 0$. In this case,

to find the second derivative is a very lengthy process. Fortunately it is not necessary, because $f \geq 0$ for all x and for $x = 0, f = 0$ therefore $x = 0$ must be a minimum.

(d) If $f(x) = \sin^2 x - \sqrt{3}\cos x$, then $f'(x) = 2\sin x \cos x + \sqrt{3}\sin x$ and this is zero wherever $\sin x = 0$ or $\cos x = -\dfrac{\sqrt{3}}{2}$ which implies in the range $0 \leq x \leq \pi$ that $x = 0, \dfrac{5\pi}{6}, \pi$. $f'(x) = \sin 2x + \sqrt{3}\sin x$, so $f''(x) = 2\cos 2x + \sqrt{3}\cos x$, which takes the values $2 + \sqrt{3}, -\frac{1}{2}, 2 - \sqrt{3}$ respectively at these three values of x. We conclude therefore that they are locations of a minimum, maximum, minimum respectively. Note how we always specify a range of values when periodic functions are involved. The values of $f(x)$ are $f(0) = -\sqrt{3}, f(5\pi/6) = 2, f(\pi) = \sqrt{3}$.

The graphs of these functions should be drawn on a graphics calculator (or computer screen) to confirm these findings.

Example 1.13 The function $f(x)$ is a continuous function on the interval $[a, b]$. Suggest an interpretation for $\lim\limits_{n\to\infty}\left\{\dfrac{(b-a)}{n}\sum\limits_{k=1}^{n} f\left(a + \dfrac{k(b-a)}{n}\right)\right\}$, and use it to evaluate $\lim\limits_{n\to\infty}\left\{\sum\limits_{k=1}^{n}\left(\dfrac{1}{n+k}\right)\right\}$.

Solution A more helpful way of writing the expression $\left\{\dfrac{(b-a)}{n}\sum\limits_{k=1}^{n} f\left(a + \dfrac{k(b-a)}{n}\right)\right\}$ is $\sum\limits_{k=1}^{n} f(\xi_k)\Delta x_k$, where $\xi_k = a + \dfrac{k}{n}(b-a)$ and $\Delta x_k = \dfrac{b-a}{n}$ (which is actually a constant with no dependence on k). This sum can then be seen to represent the sum of the areas of a number of strips, shown in Figure 1.7. Taking the limit as $n \to \infty$, that is calculating $\lim\limits_{n\to\infty}\sum\limits_{k=1}^{n}(f(\xi_k)\Delta x_k)$, means that the widths of each strip gets narrower and narrower in such a way that the area of the sum more and more closely approximates the actual area under the curve. This area is, of course, just the integral $\int_a^b f(x)\mathrm{d}x$.

Hence, $\lim\limits_{n\to\infty}\left\{\dfrac{(b-a)}{n}\sum\limits_{k=1}^{n} f\left(a + \dfrac{k(b-a)}{n}\right)\right\} = \int_a^b f(x)\mathrm{d}x$.

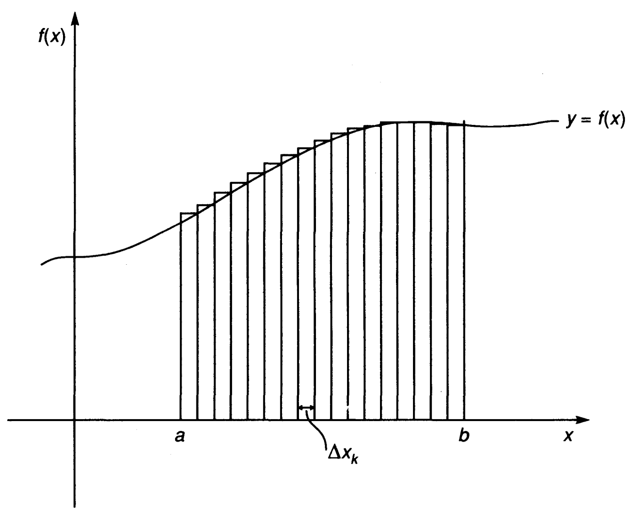


Figure 1.7 The area of the strips is approximately the integral $\int_a^b f(x)\mathrm{d}x$ for small Δx_k.

The sum $\sum\limits_{k=1}^{n}\left(\dfrac{1}{n+k}\right)$ can be re-written $\dfrac{1}{n}\sum\limits_{k=1}^{n}\left(\dfrac{1}{\left(1+\dfrac{k}{n}\right)}\right)$. Identifying $\dfrac{1}{n}$ as $\dfrac{b-a}{n}$ and $\dfrac{1}{1+\dfrac{k}{n}}$ as $f\left(a + \dfrac{k(b-a)}{n}\right)$ we conclude that $b = 1$, $a = 0$ and $f(x) = \dfrac{1}{1+x}$. Whence, $\int_a^b f(x)\mathrm{d}x = \int_0^1 \dfrac{1}{1+x}\mathrm{d}x = [\ln(1+x)]_0^1 = \ln 2$, and we conclude that the value of the limit is ln 2.

Example 1.14 Using the result $\sum_{k=1}^{n} k^2 = \frac{1}{6}n(n+1)(2n+1)$ evaluate the integral $\int_0^1 x^2 \mathrm{d}x$ from first principles.

Solution Using the limit in Example 1.12, with $f\left(\frac{k}{n}\right) = \frac{k^2}{n^2}$, $a = 0$, $b = 1$, gives

$$\int_0^1 x^2\mathrm{d}x = \lim_{n\to\infty}\left(\frac{1}{n}\sum_{k=1}^{n}\frac{k^2}{n^2}\right)$$
$$= \lim_{n\to\infty}\left(\frac{1}{n}\frac{n(n+1)(2n+1)}{6n^2}\right)$$
$$= \lim_{n\to\infty}\left(\frac{1}{6}\left(1+\frac{1}{n}\right)\left(2+\frac{1}{n}\right)\right) = \frac{1}{3}$$

Given that if $f(x) = x^n$, $f'(x) = nx^{n-1}$ (see Example 1.5), the fundamental theorem of the calculus gives $\int x^n \mathrm{d}x = \frac{x^{n+1}}{n+1} + C$. Letting $n = 2$ and inserting the limits yields $\int_0^1 x^2\mathrm{d}x = \left[\frac{x^3}{3}\right]_0^1 = \frac{1}{3}$, which agrees with our result above.

Example 1.15 Establish the integration by parts formula $\int u\mathrm{d}v = uv - \int v\mathrm{d}u$ and use it to evaluate the following indefinite integrals:

(a) $\int x\sin x\mathrm{d}x$;

(b) $\int \ln x\mathrm{d}x$;

(c) $\int x\sin^{-1}x\mathrm{d}x$.

Solution Recalling the rule for differentiation of the product of two functions (see Example 1.6): $\frac{\mathrm{d}}{\mathrm{d}x}(uv) = u\frac{\mathrm{d}v}{\mathrm{d}x} + v\frac{\mathrm{d}u}{\mathrm{d}x}$, the integration by parts formula is simply derived by integrating both sides of this equation with respect to x and rearranging. Whence, $uv = \int u\frac{\mathrm{d}v}{\mathrm{d}x}\mathrm{d}x + \int v\frac{\mathrm{d}u}{\mathrm{d}x}\mathrm{d}x$ on integration, and cancelling the differential $\mathrm{d}x$ and rearranging gives $\int u\mathrm{d}v = uv - \int v\mathrm{d}u$ as required.

(a) $\int x\sin x\mathrm{d}x$, let $x = u$, $\sin x\mathrm{d}x = \mathrm{d}v$ then $\mathrm{d}x = \mathrm{d}u$ and $-\cos x = v$. Applying the integration by parts formula gives

$$\int x\sin x\mathrm{d}x = -x\cos x + \int \cos x\mathrm{d}x = -x\cos x + \sin x + C$$

(b) $\int \ln x\mathrm{d}x$, let $\ln x = u$, $\mathrm{d}x = \mathrm{d}v$ then $\frac{\mathrm{d}x}{x} = \mathrm{d}u$, $x = v$, and integration by parts gives

$$\int \ln x\mathrm{d}x = x\ln x - \int\frac{1}{x}x\mathrm{d}x = x\ln x - x + C$$

(c) $\int x\sin^{-1}x\mathrm{d}x$, let $\sin^{-1}x = u$, $x\mathrm{d}x = \mathrm{d}v$ then $\frac{\mathrm{d}x}{\sqrt{1-x^2}} = \mathrm{d}u$, $\frac{1}{2}x^2 = v$, and integration by parts yields $\int x\sin^{-1}x\mathrm{d}x = \frac{1}{2}x^2\sin^{-1}x - \int\frac{\frac{1}{2}x^2}{\sqrt{1-x^2}}\mathrm{d}x$. The integral on the right here is not exactly obvious to evaluate, but it succumbs to the substitution $x = \sin\theta$ (or of course Computer Algebra software) to give $\frac{1}{4}\sin^{-1}x - \frac{1}{4}x\sqrt{1-x^2}$. This gives $\int x\sin^{-1}x\mathrm{d}x = \frac{1}{4}(2x^2-1)\sin^{-1}x + \frac{1}{4}x\sqrt{1-x^2} + C$ as the final answer.

Of course, all these integrals can be evaluated by software these days, the point in evaluating them by hand here is to demonstrate the method behind integration by parts. This method will surface again later, more than once.

Example 1.16 Evaluate the infinite integral $\int_e^\infty \frac{dx}{x(\ln x)^2}$, suitably defining *en route* the infinite integral.

Solution The integral $\int_a^\infty f(x)dx = \lim_{L\to\infty}\int_a^L f(x)dx$ provided the limit exists. (The integral $\int_{-\infty}^b f(x)dx$ is defined similarly.) We thus replace the infinity by L and evaluate normally, then let $L \to \infty$. The integral $\int_e^L \frac{dx}{x(\ln x)^2}$ is evaluated by substituting $u = \ln(x)$, whence $\int_e^L \frac{dx}{x(\ln x)^2} = \left[-\frac{1}{\ln x}\right]_e^L = 1 - \frac{1}{\ln L}$.

Since $\frac{1}{\ln L} \to 0$ as $L \to \infty$, we can say that $\int_e^\infty \frac{dx}{x(\ln x)^2} = 1$.

Example 1.17 Evaluate the integrals $\int_{-1}^1 \frac{dx}{x^2}$ and $\int_{-1}^1 \frac{dx}{\sqrt{x}}$.

Solution To the unsuspecting, the first integral is simply $\left[-\frac{1}{x}\right]_{-1}^1 = -2$. However, Figure 1.8 shows the graph of this function, in particular that there is a singularity $-\frac{1}{x}$ at $x = 0$. The integral therefore is not convergent, it represents an infinite area. (Beware of some Computer Algebra systems, they may also give the answer -2 for this integral!).

On the other hand, the second integral certainly exists as an indefinite integral ($2\sqrt{x}$), but square roots of negative numbers do not exist (for real numbers). We therefore discard that part of the integral that lies between -1 and 0, the rest being simply evaluated as $\left[2\sqrt{x}\right]_0^1 = 2$. Note also that the integral is the area *above* the x-axis since $\frac{1}{\sqrt{x}}$ is defined as positive.

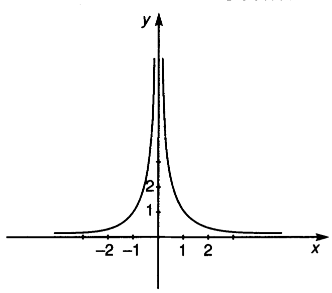


Figure 1.8 The graph of $f(x) = \frac{1}{x^2}$.

Example 1.18 Calculate the surface area and the volume formed by rotating the curve $y = \frac{1}{x}$ about the x-axis between the values $x = 1$ and $x = \infty$.

Solution At any location x, the circumference of the circle swept out by the point on the curve $y = \frac{1}{x}$ is $2\pi y$ which equals $\frac{2\pi}{x}$. The surface area is thus, by the usual calculus argument, '$\int_1^\infty \frac{2\pi}{x}dx$', and this, unfortunately, is infinite (since $\int \frac{dx}{x} = \ln x + C$). On the other hand, the volume is $\int_1^\infty \pi y^2 dx = \int_1^\infty \frac{\pi}{x^2}dx = \left[-\frac{\pi}{x}\right]_1^\infty = \pi$. The bizarre conclusion is, therefore, that a vessel so shaped could only hold a finite volume of paint ($= \pi$ in some units), but that it would take an infinite amount to paint the outside or *inside* surface of the vessel! In reality, of course, for some large value of x the diameter of the vessel $\frac{2}{x}$ is so small as to be of molecular dimensions. Mathematical models of painting vessels of such a shape must therefore involve the curve $y = \frac{1}{x}$ between $x = 1$ and $x = L$ for some $L > 1$, but not very large (perhaps 4 or 5).

Example 1.19 Calculate the distance of the centre of mass of a right circular cone, base radius a, height h, from the base.

Solution This is a standard application of integration which preludes more general applications using double and triple integrals that come later (see Chapter 9). In order to find the centre of mass, the following formula from mechanics is required:

$$\bar{x} = \frac{\int_0^h x\pi y^2 \mathrm{d}x}{\int_0^h \pi y^2 \mathrm{d}x}$$

This formula is the one-dimensional version of the more general 'first moment' formulae found not only in mechanics but also in statistics. They are of the general form:

$$\text{'mean'} = \frac{\int_{\text{vol}} \text{length} \times \mathrm{d(volume)}}{\text{(volume)}}$$

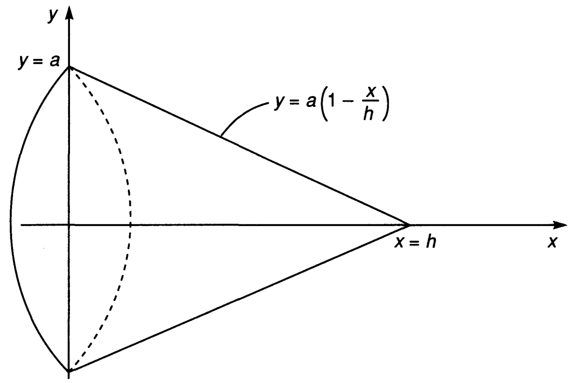


Figure 1.9 The cone of base radius a and height h.

Figure 1.9 is helpful in defining the general layout of the problem. From this figure, $y = a\left(1 - \dfrac{x}{h}\right)$ is the equation of the slanted side of the cone, so the integrals are as follows:

$$\begin{aligned}
\int_0^h \pi y^2 \mathrm{d}x &= \int_0^h \pi a^2\left(1 - \frac{x}{h}\right)^2 \mathrm{d}x \\
&= \int_0^h \pi a^2\left(1 - \frac{2x}{h} + \frac{x^2}{h^2}\right) \mathrm{d}x \\
&= \pi a^2\left[x - \frac{x^2}{h} + \frac{x^3}{3h^2}\right]_0^h = \tfrac{1}{3}\pi a^2 h \\
\int_0^h x\pi y^2 \mathrm{d}x &= \int_0^h \pi a^2 x\left(1 - \frac{2x}{h} + \frac{x^2}{h^2}\right) \mathrm{d}x \\
&= \pi a^2\left[\frac{x^2}{2} - \frac{2x^3}{3h} + \frac{x^4}{4h^2}\right]_0^h = \frac{\pi a^2 h^2}{12}
\end{aligned}$$

Dividing these two integrals gives $\bar{x} = \frac{1}{4}h$.

1.3 Exercises

1.1. Determine the values of the following limits:

(a) $\lim_{x\to 2} (x^2 + 3x + 1)$,

(b) $\lim_{x\to -4} \left\{\dfrac{2x + 8}{x^2 + x - 12}\right\}$,

(c) $\lim_{t\to 1} \left\{\dfrac{t^3 + t^2 - 5t + 3}{t^3 - 3t + 2}\right\}$,

(d) $\lim_{x\to \infty} \left\{\dfrac{\sqrt{5x^2 - 2}}{x + 3}\right\}$,

(e) $\lim_{x\to \infty} \left\{\sqrt{x^2 + bx} - \sqrt{x^2 + ax}\right\}$,

(f) $\lim_{x\to 0} \left\{\dfrac{\sin 3x}{\sin 5x}\right\}$,

(g) $\lim_{\theta\to 0} \left\{\dfrac{\sin 7\theta}{\theta}\right\}$,

(h) $\lim_{h\to 0} \left\{\dfrac{h}{\tanh}\right\}$,

(i) $\lim_{t\to 0} \left\{\dfrac{t^2}{1 - \cos^2 t}\right\}$,

(j) $\lim_{x\to \frac{\pi}{4}} \left\{\dfrac{\tan x - 1}{x - \pi/4}\right\}$.

1.2. Supplies are dropped by a parachute, the speed of which is given by the formula $v = 8[1 - (0.09)^t]$ where t is time. Calculate the speed that is approached as t gets very large (the terminal velocity). How long before v reaches 95 per cent of this speed?

1.3. Establish **from first principles** that:
(a) $\frac{d}{dx}(x^2) = 2x$,
(b) $\frac{d}{dx}(\cos x) = -\sin x$,
(c) $\frac{d}{dx}(e^x) = e^x$,
(d) $\frac{d}{dx}\left(\frac{1}{x}\right) = -\frac{1}{x^2}$.
(For part (c) you may assume that $e^t \approx 1 + t$ for small t.)

1.4. Establish using limits that

$$\frac{d}{dx}(y^n) = ny^{n-1}\frac{dy}{dx} \text{ (implicit differentiation).}$$

1.5. Determine whether the following function has a derivative at $x = 0$:

$$f(x) = \begin{cases} 1 - \cos x & x < 0 \\ x^2 & x \geq 0 \end{cases}$$

Does it have a *second* derivative at $x = 0$?

1.6. Find the points at which the following functions have extreme values, and classify them as maxima, minima or points of inflection:
(a) $x^2 - 5x + 6$,
(b) $x^4 - 6x^2 - 3$,
(c) $x\tan x$ $(-\pi/2 < x < \pi/2)$
(d) $\frac{x}{x^2 + 2}$.

1.7. [Do not use either Rolle's Theorem or the Mean Value Theorem for this question].
Why does Rolle's Theorem not work for the function $y = \tan x$ even though $\tan 0 = \tan\pi = 0$?

1.8. Demonstrate that if $f(x)$ is continuous on $[a, b]$ and differentiable on (a, b), and if $f(a) = f(b)$, then there is a point c in (a, b) where $f'(c) = 0$ (see Example 1.8).

1.9. An object moves in a straight line such that its position $x = x(t)$ where x is a differentiable function. Interpret Rolle's Theorem and the Mean Value Theorem in terms of the motion of the object.

1.10. Derive the first three terms of the Taylor Series about $x = 0$ (sometimes called Maclaurin Series) for the following functions:
(a) $\sin x$,
(b) $\cos x$,
(c) $\ln(1 + x)$,
(d) $\sqrt{1 + x}$.
Determine the range of x for which each series is valid.

1.11. A polynomial is its own Maclaurin Series. Explain this statement.

1.12. Expand each of the following functions in the terms indicated:
(a) $x\ln x$ in powers of $(x - 2)$,
(b) $(1 - 2x)^{-3}$ in powers of $(x + 2)$,
(c) $\sin x$ in powers of $(x - \pi)$,
in each case finding at least three terms.

1.13. Choose an appropriate Taylor Series to approximate $\sqrt{61}$ to three decimal places of accuracy.

1.14. Use L'Hôpital's Rule to determine the values of the following limits:
(a) $\lim_{x\to 0}\left\{\frac{1 - \cos 4x}{9x^2}\right\}$,
(b) $\lim_{x\to 0}\left\{\frac{\sec x - 1}{x\sec x}\right\}$,
(c) $\lim_{x\to 0}\{x^2(1 + \cot^2 3x)\}$,
(d) $\lim_{x\to \pi}\left\{\frac{\sin x}{x - \pi}\right\}$,
(e) $\lim_{x\to 0^+}\left\{\frac{\sin x}{5\sqrt{x}}\right\}$,
(f) $\lim_{t\to 0}\left\{\frac{1 - \cos 5t}{\cos 7t - 1}\right\}$,
(g) $\lim_{x\to 0}\left\{\frac{e^{ax} - e^{bx}}{x}\right\}$,
(h) $\lim_{x\to 0}\left\{\frac{a^x - 1}{x}\right\}$ $(a > 0)$,
(i) $\lim_{\theta\to 0}\left\{\frac{\sin^2\theta - \sin(\theta^2)}{\theta^4}\right\}$,
(j) $\lim_{x\to 1}\left\{\sqrt{\frac{\ln x}{x^4 - 1}}\right\}$,
(k) $\lim_{x\to 0^+}\{\tan x\ln x\}$,
(l) $\lim_{x\to 1} x^{1/(1-x)}$,
(m) $\lim_{x\to\infty}(\ln x)^{1/x}$,
(n) $\lim_{x\to\infty}(3^x + 5^x)^{1/x}$.

1.15. Investigate the limit $\lim_{x\to 0}\left\{\frac{\sin(\tan x) - \tan(\sin x)}{x^7}\right\}$ (you will need access to a graphics calculator or computer software).

1.16. Use the Newton–Raphson method to find $\sqrt{2}$, $\sqrt{7}$ and $\sqrt[3]{6}$ to one more place than is contained on your calculator.

1.17. Apply the Newton–Raphson method to the equation $x^2 = a$ to derive the 'well known' rule

$$x_{n+1} = \tfrac{1}{2}\left(x_n + \frac{a}{x_n}\right), \quad n = 1, 2, \ldots$$

for finding the square root of a number a.

1.18. Use the limit in Example 1.12 to evaluate the integral $\int_0^1 x^3 dx$ from first principles. (You will need the sum $\sum_{k=1}^{n} k^3 = \frac{n^2}{4}(n + 1)^2$.)

1.19. Evaluate the following integrals using integration by parts:
(a) $\int_0^1 \tan^{-1}x\,dx$,
(b) $\int x\ln x\,dx$,
(c) $\int \frac{xe^x}{(x + 1)^2}dx$.

1.20. If $I_n = \int\cos^n x\,dx$ show that $I_n = \frac{1}{n}\cos^{n-1}x\sin x + \frac{n - 1}{n}I_{n-2}$ and hence evaluate $\int\cos^4 x\,dx$ and $\int\cos^5 x\,dx$.

1.21. Using a similar formula to that derived in the last question (called a *reduction formula*) evaluate $\int \sin^4 x\mathrm{d}x$ and $\int \sin^5 x\mathrm{d}x$.

1.22. Evaluate the following improper integrals:

(a) $\displaystyle\int_1^\infty \frac{\mathrm{d}x}{x^3}$,

(b) $\displaystyle\int_4^\infty \frac{2\mathrm{d}x}{x^2 - 1}$,

(c) $\displaystyle\int_0^\infty x\mathrm{e}^{-x^2}\mathrm{d}x$,

(d) $\displaystyle\int_1^\infty \frac{\mathrm{d}x}{\sqrt{x}}$,

(e) $\displaystyle\int_0^\infty \frac{\mathrm{d}x}{a^2 + b^2x^2}$.

1.23. Calculate the surface area and the volume formed by rotating the curve $y = 4 - x^2$ once around the x-axis between the values $x = 0$ and $x = 2$.

1.24. A water clock or clepsydra is formed by rotating the curve $y = bx^4$ about the y-axis. Calculate the height of its centre of mass if the vessel is of total height h.

Topic Guide

Definition
Chain Rule
Jacobians
Euler's Theorem

2 Partial Differentiation

2.1 Fact Sheet

Chapter 1 deals with functions of a single variable, written $f(x)$. We now turn to functions of many variables, in general written $f(x_1, x_2, \ldots\ldots, x_n)$ where $x_1, x_2, \ldots, x_n$ are n independent variables, but most attention will be focused on the special (and simplest) case, $n = 2$, where the function is written $f(x, y)$. This two-variable calculus has the most applications to the real world and has the merit of being applied to problems that can be visualised in three-dimensional space. The general scheme is therefore to state results in general, but to use two dimensions when doing examples, except in rare circumstances.

Partial differentiation is denoted by $\frac{\partial f}{\partial x}$ rather than $\frac{df}{dx}$, and has the definition $\frac{\partial f}{\partial x} = \lim_{h \to 0}\left(\frac{f(x+h, y) - f(x, y)}{h}\right)$, and $\frac{\partial f}{\partial y} = \lim_{k \to 0}\left(\frac{f(x, y+k) - f(x, y)}{k}\right)$. The symbol ∂ indicates that all variables other than that following this symbol in the denominator are to be treated as constants. Functions of n variables will thus have n first-order partial derivatives. The product and quotient rules of differentiation generalise straightforwardly, in that a product or quotient is evaluated as in the single-variable case, all other variables except the one being differentiated simply assuming the role of constants. The function of a function rule, however, is replaced by the *chain rule*. Let us state this in general, assuming that f is a function of the n variables $(x_1, x_2, \ldots, x_n)$ and that each of these variables, in turn, is a function of m more variables $(u_1, u_2, \ldots, u_m)$. The following equations govern the transformation from the x variables to the u variables and generalise the formula $\frac{df}{dx} = \frac{df}{du}\frac{du}{dx}$.

$$
\begin{aligned}
\frac{\partial f}{\partial u_1} &= \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial u_1} + \frac{\partial f}{\partial x_2}\frac{\partial x_2}{\partial u_1} + \ldots + \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial u_1} \\
\frac{\partial f}{\partial u_2} &= \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial u_2} + \frac{\partial f}{\partial x_2}\frac{\partial x_2}{\partial u_2} + \ldots + \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial u_2} \\
&\vdots \\
\frac{\partial f}{\partial u_m} &= \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial u_m} + \frac{\partial f}{\partial x_2}\frac{\partial x_2}{\partial u_m} + \ldots + \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial u_m}
\end{aligned}
$$

There is a similar array of formulae for obtaining partial derivatives of f with respect to the x variables in terms of the u variables:

$$\frac{\partial f}{\partial x_1} = \frac{\partial f}{\partial u_1}\frac{\partial u_1}{\partial x_1} + \frac{\partial f}{\partial u_2}\frac{\partial u_2}{\partial x_1} + \ldots + \frac{\partial f}{\partial u_m}\frac{\partial u_m}{\partial x_1}$$

$$\frac{\partial f}{\partial x_2} = \frac{\partial f}{\partial u_1}\frac{\partial u_1}{\partial x_2} + \frac{\partial f}{\partial u_2}\frac{\partial u_2}{\partial x_2} + \ldots + \frac{\partial f}{\partial u_m}\frac{\partial u_m}{\partial x_2}$$

$$\vdots$$

$$\frac{\partial f}{\partial x_n} = \frac{\partial f}{\partial u_1}\frac{\partial u_1}{\partial x_n} + \frac{\partial f}{\partial u_2}\frac{\partial u_2}{\partial x_n} + \ldots + \frac{\partial f}{\partial u_m}\frac{\partial u_m}{\partial x_n}$$

It is worth noting here that there is no relationship between $\frac{\partial u_i}{\partial x_j}$ and $\frac{\partial x_j}{\partial u_i}$ analogous to $\frac{\mathrm{d}y}{\mathrm{d}x} = \left(\frac{\mathrm{d}x}{\mathrm{d}y}\right)^{-1}$. A notation that really comes into its own when dealing with partial derivatives is the suffix derivative notation whereby $\frac{\partial f}{\partial x}$ is written f_x and $\frac{\partial f}{\partial y}$ is written f_y. If higher derivatives are required, then the suffix is repeated or added to. For example, $\frac{\partial^2 f}{\partial x^2} = f_{xx}$ and $\frac{\partial^3 f}{\partial x \partial y^2} = f_{xyy}$ (etc.). Another useful concept is that of a *differential*, $\mathrm{d}x$ (say) which actually has the value zero, but occurs in integrals, and obeys rules such as if $x = x(u_1, u_2)$ then $\mathrm{d}x = \frac{\partial x}{\partial u_1}\mathrm{d}u_1 + \frac{\partial x}{\partial u_2}\mathrm{d}u_2$.

The *Jacobian* is defined in terms of two variables as follows:

$$J(u_1, u_2) = \begin{vmatrix} \frac{\partial x_1}{\partial u_1} & \frac{\partial x_2}{\partial u_1} \\ \frac{\partial x_1}{\partial u_2} & \frac{\partial x_2}{\partial u_2} \end{vmatrix} = \frac{\partial x_1}{\partial u_1}\frac{\partial x_2}{\partial u_2} - \frac{\partial x_1}{\partial u_2}\frac{\partial x_2}{\partial u_1}$$

(For those who know nothing of determinants, there is a short résumé in Chapter 4.) The more general definition is derived straightforwardly:

$$J(u_1, u_2, \ldots, u_n) = \begin{vmatrix} \frac{\partial x_1}{\partial u_1} & \frac{\partial x_2}{\partial u_1} & \cdots & \frac{\partial x_n}{\partial u_1} \\ \frac{\partial x_1}{\partial u_2} & \frac{\partial x_2}{\partial u_2} & \cdots & \frac{\partial x_n}{\partial u_2} \\ & \vdots & & \\ \frac{\partial x_1}{\partial u_n} & \frac{\partial x_2}{\partial u_n} & \cdots & \frac{\partial x_n}{\partial u_n} \end{vmatrix}$$

The definition of $J(x_1, x_2, \ldots, x_n)$ is as above but with the roles of x and u reversed. It can be shown that $J(x_1, x_2, \ldots, x_n) = (J(u_1, u_2, \ldots, u_n))^{-1}$. Note that there are as many x's as u's, otherwise the determinant could not be defined. There are useful results that stem from familiarity with Jacobians. For example: if $f(x, y)$ and $g(x, y)$ are functions of two variables such that $u = f(x, y)$ and $v = g(x, y)$, and further $\frac{\partial(u, v)}{\partial(x, y)} = 0$, then there is a functional relationship between u and v, $\phi(u, v) = 0$ (see Example 2.4).

There are many results and theorems in the field of analysis involving partial derivatives, one of the more useful of which is due to the prolific Swiss mathematician Leonhard Euler and is called *Euler's Theorem for Homogeneous Functions.* A function $f(x_1, x_2, \ldots, x_n)$ is said to be homogeneous of degree p if

$$f(tx_1, tx_2, \ldots, tx_n) \equiv t^p f(x_1, x_2, \ldots, x_n)$$

If $f(x_1, x_2, \ldots, x_n)$ is a homogeneous function of degree p then Euler's Theorem states that

$$x_1 \frac{\partial f}{\partial x_1} + x_2 \frac{\partial f}{\partial x_2} + \ldots + x_n \frac{\partial f}{\partial x_n} = pf$$

There are further, more complicated theorems on these lines (also due to Euler) involving higher derivatives, but these are outside the scope of this book.

Another very useful result is *Leibniz's Rule*, sometimes given the name *Differentiation under the Integral Sign*, which goes as follows:

$$\frac{\partial}{\partial \alpha}\left(\int_{a(\alpha)}^{b(\alpha)} f(x, \alpha)\mathrm{d}x\right) = \int_{a(\alpha)}^{b(\alpha)} \frac{\partial f}{\partial \alpha}\mathrm{d}x + \frac{\mathrm{d}b}{\mathrm{d}\alpha} f(x, \alpha)\big|_{x=b(\alpha)} - \frac{\mathrm{d}a}{\mathrm{d}\alpha} f(x, \alpha)\big|_{x=a(\alpha)}$$

A useful concept is the *total differential*. For a function f of n variables $f(x_1, x_2, \ldots, x_n)$ the total differential is given by the expression

$$\mathrm{d}f = \frac{\partial f}{\partial x_1}\mathrm{d}x_1 + \frac{\partial f}{\partial x_2}\mathrm{d}x_2 + \ldots + \frac{\partial f}{\partial x_n}\mathrm{d}x_n$$

which follows from the application of the chain rule. Its usefulness lies in being able to determine the error in f due to the errors in the x_i ($i = 1, 2, \ldots, n$) (see Example 2.9).

2.2 Worked Examples

A function of two variables $f(x, y)$ is defined by $x^3 + 3x^2y^2 + y^3$. Find the partial derivatives: f_x, f_y, f_{xx}, f_{xy} and f_{yy}.

Example 2.1

Solution Using the elementary rules of differentiation, we obtain

$$\frac{\partial f}{\partial x} = f_x = 3x^2 + 6xy^2$$

$$\frac{\partial^2 f}{\partial x^2} = f_{xx} = 6x + 6y^2$$

$$\frac{\partial f}{\partial y} = f_y = 6x^2y + 3y^2$$

The mixed derivative $\dfrac{\partial^2 f}{\partial x \partial y} = f_{xy} = 12xy$. Note that $f_{xy} = f_{yx}$ unless the function $f(x, y)$ has discontinuous second-order partial derivatives. This is rare and usually such functions have to be purposely manufactured:

$$\frac{\partial^2 f}{\partial y^2} = f_{yy} = 6y + 6x^2$$

The expression $x^3 + 3x^2y^2 + y^3$ is symmetric, that is it remains the same if x and y swap roles. This fact is useful as either a check on working, or as a means of finding other partial derivatives (see Example 2.2). Hence in this example, f_{yy} can be obtained directly from f_{xx} by interchanging the roles of x and y. If f is symmetric, then so is f_{xy}. Computer algebra systems can also be used to check the answers.

Example 2.2 Find *all* first and second partial derivatives of the function $f(x, y, z) = \sin(xyz)$ and show that $x^2 f_{xx} + y^2 f_{yy} + z^2 f_{zz} + 3x^2y^2z^2 f = 0$.

Solution Since $\sin(xyz)$ is symmetric in x, y and z, it is only necessary to find the three derivatives f_x, f_{xy} and f_{xx} directly. The remainder can be found indirectly by symmetry. Differentiating we obtain

$$\begin{aligned} f_x &= yz\cos(xyz) \\ f_{xy} &= z\cos(xyz) - xyz^2\sin(xyz) \\ f_{xx} &= -y^2z^2\sin(xyz) \end{aligned}$$

Using symmetry arguments, the rest of the partial derivatives of f can be written immediately as follows

$$\begin{aligned} f_y &= xz\cos(xyz) \\ f_z &= xy\cos(xyz) \\ f_{yz} &= x\cos(xyz) - x^2yz\sin(xyz) \\ f_{zx} &= y\cos(xyz) - xy^2z\sin(xyz) \\ f_{yy} &= -x^2z^2\sin(xyz) \\ f_{zz} &= -x^2y^2\sin(xyz) \end{aligned}$$

Hence, $x^2f_{xx} + y^2f_{yy} + z^2f_{zz} = -3x^2y^2z^2\sin(xyz) = -3x^2y^2z^2f$ so that $x^2f_{xx} + y^2f_{yy} + z^2f_{zz} + 3x^2y^2z^2f = 0$ as required.

Example 2.3 If $f(x, y, z) = \dfrac{z}{x^2 + y^2}$ and $x = u + v$, $y = u$, $z = uv$ find $\dfrac{\partial f}{\partial x}$, $\dfrac{\partial f}{\partial y}$, $\dfrac{\partial f}{\partial z}$, $\dfrac{\partial f}{\partial u}$ and $\dfrac{\partial f}{\partial v}$, and explain why $\dfrac{\partial f}{\partial y} \neq \dfrac{\partial f}{\partial u}$ even though $y = u$.

Solution

$$\frac{\partial f}{\partial x} = -\frac{z}{(x^2 + y^2)^2} \,.\, 2x = -\frac{2xz}{(x^2 + y^2)^2}$$

$$\frac{\partial f}{\partial y} = -\frac{2yz}{(x^2 + y^2)^2}, \quad \frac{\partial f}{\partial z} = \frac{1}{x^2 + y^2}$$

In order to find the other two partial derivatives, we could either substitute directly or use the chain rule.

$$\begin{aligned} \frac{\partial f}{\partial u} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial u} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial u} \\ &= \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} + v\frac{\partial f}{\partial z} \\ &= -\frac{2xz}{(x^2 + y^2)^2} - \frac{2yz}{(x^2 + y^2)^2} + \frac{v}{x^2 + y^2} \end{aligned}$$

So

$$\begin{aligned} \frac{\partial f}{\partial u} &= -2f\left(\frac{x + y}{x^2 + y^2}\right) + \frac{v}{x^2 + y^2} \\ &= -2f^2\,\frac{x + y}{z} + vf\,\frac{1}{z} \\ &= -2f^2\,\frac{(2u + v)}{uv} + \frac{f}{u} \end{aligned}$$

and

$$\begin{aligned} \frac{\partial f}{\partial v} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial v} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial v} \\ &= \frac{\partial f}{\partial x} + u\frac{\partial f}{\partial z} \end{aligned}$$

$$= -\frac{2xz}{(x^2+y^2)^2} + \frac{u}{x^2+y^2}$$

$$= -2f^2\frac{(u+v)}{uv} + \frac{f}{v}$$

Obviously $\frac{\partial f}{\partial y} \neq \frac{\partial f}{\partial u}$; this is because although $y = u$, different variables are being held constant (x and z in the case of $\frac{\partial f}{\partial y}$, v in the case of $\frac{\partial f}{\partial u}$).

Example 2.4 If $x = r\cos\theta$, $y = r\sin\theta$ show that for a function $f(x,y)$ which has continuous partial derivatives:

$$\frac{\partial f}{\partial x} = \cos\theta\frac{\partial f}{\partial r} - \frac{\sin\theta}{r}\frac{\partial f}{\partial \theta}$$

and

$$\frac{\partial f}{\partial y} = \sin\theta\frac{\partial f}{\partial r} + \frac{\cos\theta}{r}\frac{\partial f}{\partial \theta}$$

Solution The chain rule states that

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial r}\frac{\partial r}{\partial x} + \frac{\partial f}{\partial \theta}\frac{\partial \theta}{\partial x}$$

and

$$\frac{\partial f}{\partial y} = \frac{\partial f}{\partial r}\frac{\partial r}{\partial y} + \frac{\partial f}{\partial \theta}\frac{\partial \theta}{\partial y}$$

We are given that $x = r\cos\theta$, $y = r\sin\theta$ so $\tan\theta = \frac{y}{x}$ and $r^2 = x^2 + y^2$.
Differentiating these we get
$\sec^2\theta\frac{\partial \theta}{\partial x} = -\frac{y}{x^2}$, $\sec^2\theta\frac{\partial \theta}{\partial y} = \frac{1}{x}$, $2r\frac{\partial r}{\partial x} = 2x$, $2r\frac{\partial r}{\partial y} = 2y$.
So $\frac{\partial \theta}{\partial x} = -\frac{y}{x^2\sec^2\theta} = -\frac{r\sin\theta}{r^2} = -\frac{\sin\theta}{r}$, $\frac{\partial \theta}{\partial y} = \frac{1}{x\sec^2\theta} = \frac{\cos\theta}{r}$,
and $\frac{\partial r}{\partial x} = \frac{x}{r} = \cos\theta$, $\frac{\partial r}{\partial y} = \frac{y}{r} = \sin\theta$.
From this we find

$$\frac{\partial f}{\partial x} = \cos\theta\frac{\partial f}{\partial r} - \frac{\sin\theta}{r}\frac{\partial f}{\partial \theta}$$

and

$$\frac{\partial f}{\partial y} = \sin\theta\frac{\partial f}{\partial r} + \frac{\cos\theta}{r}\frac{\partial f}{\partial \theta}$$

as required.

Example 2.5 A variable ϕ obeys Laplace's equation in two-dimensional Cartesian co-ordinates: $\frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} = 0$.
If $x = r\cos\theta$, $y = r\sin\theta$, find the (r, θ) equation for ϕ.

Solution First of all, we need to find the first derivatives $\frac{\partial \phi}{\partial x}$ and $\frac{\partial \phi}{\partial y}$ in terms of r and θ:

$$\frac{\partial \phi}{\partial x} = \frac{\partial \phi}{\partial r}\frac{\partial r}{\partial x} + \frac{\partial \phi}{\partial \theta}\frac{\partial \theta}{\partial x}$$

$$= \cos\theta\frac{\partial \phi}{\partial r} - \frac{\sin\theta}{r}\frac{\partial \phi}{\partial \theta}$$

$$\frac{\partial \phi}{\partial y} = \frac{\partial \phi}{\partial r}\frac{\partial r}{\partial y} + \frac{\partial \phi}{\partial \theta}\frac{\partial \theta}{\partial y}$$

$$= \sin\theta\frac{\partial \phi}{\partial r} + \frac{\cos\theta}{r}\frac{\partial \phi}{\partial \theta}$$

Taking these as operator identities gives

$$\begin{aligned}
\frac{\partial^2\phi}{\partial x^2} &= \left(\cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial\theta}\right)\left(\cos\theta\frac{\partial\phi}{\partial r} - \frac{\sin\theta}{r}\frac{\partial\phi}{\partial\theta}\right)\\
&= \cos\theta\frac{\partial}{\partial r}\left(\cos\theta\frac{\partial\phi}{\partial r} - \frac{\sin\theta}{r}\frac{\partial\phi}{\partial\theta}\right) - \frac{\sin\theta}{r}\frac{\partial}{\partial\theta}\left(\cos\theta\frac{\partial\phi}{\partial r} - \frac{\sin\theta}{r}\frac{\partial\phi}{\partial\theta}\right)\\
&= \cos^2\theta\frac{\partial^2\phi}{\partial r^2} + \frac{\cos\theta\sin\theta}{r^2}\frac{\partial\phi}{\partial\theta} - \frac{\cos\theta\sin\theta}{r}\frac{\partial^2\phi}{\partial r\partial\theta}\\
&\quad - \frac{\sin\theta}{r}\left(-\sin\theta\frac{\partial\phi}{\partial r} + \cos\theta\frac{\partial^2\phi}{\partial r\partial\theta} - \frac{\cos\theta}{r}\frac{\partial\phi}{\partial\theta} - \frac{\sin\theta}{r}\frac{\partial^2\phi}{\partial\theta^2}\right)\\
&= \cos^2\theta\frac{\partial^2\phi}{\partial r^2} + \frac{2\cos\theta\sin\theta}{r^2}\frac{\partial\phi}{\partial\theta} - \frac{2\cos\theta\sin\theta}{r}\frac{\partial^2\phi}{\partial r\partial\theta} + \frac{\sin^2\theta}{r}\frac{\partial\phi}{\partial r} + \frac{\sin^2\theta}{r^2}\frac{\partial^2\phi}{\partial\theta^2}
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial^2\phi}{\partial y^2} &= \left(\sin\theta\frac{\partial}{\partial r} + \frac{\cos\theta}{r}\frac{\partial}{\partial\theta}\right)\left(\sin\theta\frac{\partial\phi}{\partial r} + \frac{\cos\theta}{r}\frac{\partial\phi}{\partial\theta}\right)\\
&= \sin\theta\frac{\partial}{\partial r}\left(\sin\theta\frac{\partial\phi}{\partial r} + \frac{\cos\theta}{r}\frac{\partial\phi}{\partial\theta}\right) + \frac{\cos\theta}{r}\left(\sin\theta\frac{\partial\phi}{\partial r} + \frac{\cos\theta}{r}\frac{\partial\phi}{\partial\theta}\right)\\
&= \sin^2\theta\frac{\partial^2\phi}{\partial r^2} - \frac{\cos\theta\sin\theta}{r^2}\frac{\partial\phi}{\partial\theta} + \frac{\cos\theta\sin\theta}{r}\frac{\partial^2\phi}{\partial r\partial\theta}\\
&\quad + \frac{\cos\theta}{r}\left(\cos\theta\frac{\partial\phi}{\partial r} + \sin\theta\frac{\partial^2\phi}{\partial r\partial\theta} - \frac{\sin\theta}{r}\frac{\partial\phi}{\partial\theta} + \frac{\cos\theta}{r}\frac{\partial^2\phi}{\partial\theta^2}\right)\\
&= \sin^2\theta\frac{\partial^2\phi}{\partial r^2} - \frac{2\cos\theta\sin\theta}{r^2}\frac{\partial\phi}{\partial\theta} + \frac{2\cos\theta\sin\theta}{r}\frac{\partial^2\phi}{\partial r\partial\theta} + \frac{\cos^2\theta}{r}\frac{\partial\phi}{\partial r} + \frac{\cos^2\theta}{r^2}\frac{\partial^2\phi}{\partial\theta^2}
\end{aligned}$$

Adding together (and cancelling terms) gives

$$\frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} = \frac{\partial^2\phi}{\partial r^2} + \frac{1}{r}\frac{\partial\phi}{\partial r} + \frac{1}{r^2}\frac{\partial^2\phi}{\partial\theta^2} = 0$$

as the plane polar co-ordinate version of Laplace's equation. This equation is written in co-ordinate free notation as $\nabla^2\phi = 0$ (see Chapter 7).

Example 2.6 The ideal gas law states that $pV = mRT$ where p is the gas pressure, V is the volume of gas, T is the temperature of the gas, and m and R are constants. Show that $\dfrac{\partial T}{\partial p}\cdot\dfrac{\partial p}{\partial V}\cdot\dfrac{\partial V}{\partial T} = -1$.

Solution Writing the ideal gas law in three different ways to facilitate partial differentiation, $p = \dfrac{mRT}{V}$, $T = \dfrac{pV}{mR}$ and $V = \dfrac{mRT}{p}$. Thus, $\dfrac{\partial p}{\partial V} = -\dfrac{mRT}{V^2}$, $\dfrac{\partial T}{\partial p} = \dfrac{V}{mR}$ and $\dfrac{\partial V}{\partial T} = \dfrac{mR}{p}$. Multiplying these three expressions together gives $\dfrac{\partial T}{\partial p}\cdot\dfrac{\partial p}{\partial V}\cdot\dfrac{\partial V}{\partial T} = \dfrac{V}{mR}\cdot\left(-\dfrac{mRT}{V^2}\right)\cdot\dfrac{mR}{p} = -\dfrac{mRT}{pV} = -1$ as required.

Example 2.7 If $\psi = (x + y)\phi\left(\dfrac{y}{x}\right)$ where ϕ is an arbitrary function with continuous second-order derivatives, show that

(*a*) $x\dfrac{\partial\psi}{\partial x} + y\dfrac{\partial\psi}{\partial y} = \psi$

(*b*) $x^2\dfrac{\partial^2\psi}{\partial x^2} + 2xy\dfrac{\partial^2\psi}{\partial x\partial y} + y^2\dfrac{\partial^2\psi}{\partial y^2} = 0$

Solution This kind of problem is an exercise in the manipulation of partial derivatives; its novelty lies in the use of operator techniques.

(a) Given $\psi = (x + y)\phi\left(\frac{y}{x}\right)$

$$\frac{\partial\psi}{\partial x} = \phi\left(\frac{y}{x}\right) + (x + y)\phi'\left(\frac{y}{x}\right)\left(-\frac{y}{x^2}\right)$$

$$\frac{\partial\psi}{\partial y} = \phi\left(\frac{y}{x}\right) + (x + y)\phi'\left(\frac{y}{x}\right)\left(\frac{1}{x}\right)$$

$$x\frac{\partial\psi}{\partial x} + y\frac{\partial\psi}{\partial y} = x\phi + (x + y)\left(-\frac{y}{x}\right)\phi' + (x + y)\left(\frac{y}{x}\right)\phi' + y\phi$$

$$= (x + y)\phi = \psi$$

as required. (Dash denotes derivative with respect to the argument of ϕ, y/x.)

Part (b) can most easily be proved by using operator techniques. First of all, part (a) yields $x\frac{\partial\psi}{\partial x} + y\frac{\partial\psi}{\partial y} = \psi$ so the operator $x\frac{\partial\psi}{\partial x} + y\frac{\partial\psi}{\partial y}$ is the identity operator. Therefore, operating twice on ψ gives

$$\left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right)^2 \psi = \left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right)\psi = \psi$$

$$\text{or } \left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right)\left(x\frac{\partial\psi}{\partial x} + y\frac{\partial\psi}{\partial y}\right) = \psi$$

So

$$x\frac{\partial}{\partial x}\left(x\frac{\partial\psi}{\partial x}\right) + x\frac{\partial}{\partial x}\left(y\frac{\partial\psi}{\partial y}\right) + y\frac{\partial}{\partial y}\left(x\frac{\partial\psi}{\partial x}\right) + y\frac{\partial}{\partial y}y\left(\frac{\partial\psi}{\partial y}\right) = \psi$$

$$x\frac{\partial\psi}{\partial x} + x^2\frac{\partial^2\psi}{\partial x^2} + xy\frac{\partial^2\psi}{\partial x\partial y} + xy\frac{\partial^2\psi}{\partial x\partial y} + y\frac{\partial\psi}{\partial y} + y^2\frac{\partial^2\psi}{\partial x\partial y} = \psi$$

$$x\frac{\partial\psi}{\partial x} + y\frac{\partial\psi}{\partial y} + x^2\frac{\partial^2\psi}{\partial x^2} + 2xy\frac{\partial^2\psi}{\partial x\partial y} + y^2\frac{\partial^2\psi}{\partial y^2} = \psi$$

Given, from part (a), that the first two terms are ψ, we have deduced that

$$x^2\frac{\partial^2\psi}{\partial x^2} + 2xy\frac{\partial^2\psi}{\partial x\partial y} + y^2\frac{\partial^2\psi}{\partial y^2} = 0$$

This is extraordinarily difficult to prove by other means.

Example 2.8 If $u = f(x, y)$ and $v = g(x, y)$ where f and g are suitably well behaved functions (continuously differentiable), show that there is a functional relationship of the kind $\phi(u, v) = 0$ if and only if $\frac{\partial(u, v)}{\partial(x, y)} = 0$.

Solution As is quite typical with this 'necessary and sufficient' kind of proof, the result is much easier to prove one way. Let's do the easy one first. This means we assume that $\phi(u, v) = 0$, the *necessary* condition. From this

$$\mathrm{d}\phi = \frac{\partial\phi}{\partial u}\mathrm{d}u + \frac{\partial\phi}{\partial v}\mathrm{d}v = 0$$

However we cannot equate coefficients of du and dv to zero since they are not independent and would in any case immediately imply that $\frac{\partial\phi}{\partial u} = 0$ and $\frac{\partial\phi}{\partial v} = 0$, that is ϕ is *independent* of u and v

prohibiting any functional relationship which is the initial assumption. We can progress if we write ϕ in terms of x and y so that $\mathrm{d}\phi = \dfrac{\partial \phi}{\partial x}\,\mathrm{d}x + \dfrac{\partial \phi}{\partial y}\,\mathrm{d}y = 0$. This *does* imply that $\dfrac{\partial \phi}{\partial x} = 0$ and $\dfrac{\partial \phi}{\partial y} = 0$ as $\mathrm{d}x$ and $\mathrm{d}y$ *are* independent, and writing in terms of the derivatives of u and v via the chain rule, gives the simultaneous equations

$$\frac{\partial \phi}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial \phi}{\partial v}\frac{\partial v}{\partial x} = 0$$

and

$$\frac{\partial \phi}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial \phi}{\partial v}\frac{\partial v}{\partial y} = 0$$

The condition that $\dfrac{\partial \phi}{\partial u}$ and $\dfrac{\partial \phi}{\partial v}$ are not both zero is that the determinant of the coefficients is zero, that is

$$\begin{vmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial v}{\partial x} \\ \dfrac{\partial u}{\partial y} & \dfrac{\partial v}{\partial y} \end{vmatrix} = 0$$

which is $\dfrac{\partial(u, v)}{\partial(x, y)} = 0$. This is the required Jacobian, and proves the condition is necessary. To prove *sufficiency* we start with the Jacobian being zero and deduce that there must be a functional relationship between u and v.

Since $u = f(x, y)$ then, theoretically, we can make x the subject of the formula to obtain $x = F(u, y)$. Thus $u = f(F(u, y), y)$ and $v = g(F(u, y), y)$ from which, using the chain rule

$$\mathrm{d}u = \frac{\partial u}{\partial x}\,\mathrm{d}x + \frac{\partial u}{\partial y}\,\mathrm{d}y = \frac{\partial u}{\partial x}\left(\frac{\partial F}{\partial u}\,\mathrm{d}u + \frac{\partial F}{\partial y}\,\mathrm{d}y\right) + \frac{\partial u}{\partial y}\,\mathrm{d}y$$

or

$$\mathrm{d}u = \frac{\partial u}{\partial x}\frac{\partial F}{\partial u}\,\mathrm{d}u + \left(\frac{\partial u}{\partial x}\frac{\partial F}{\partial y} + \frac{\partial u}{\partial y}\right)\mathrm{d}y$$

This gives $\quad 1 = \dfrac{\partial u}{\partial x}\dfrac{\partial F}{\partial u}$ and $\dfrac{\partial F}{\partial y} = -\dfrac{\partial u/\partial y}{\partial u/\partial x}$ provided $\dfrac{\partial u}{\partial x} \neq 0$.
Proceeding similarly with $\mathrm{d}v$ we find that

$$\mathrm{d}v = \frac{\partial v}{\partial x}\,\mathrm{d}x + \frac{\partial v}{\partial y}\,\mathrm{d}y = \frac{\partial v}{\partial x}\left(\frac{\partial F}{\partial u}\,\mathrm{d}u + \frac{\partial F}{\partial y}\,\mathrm{d}y\right) + \frac{\partial v}{\partial y}\,\mathrm{d}y$$

or

$$\mathrm{d}v = \frac{\partial v}{\partial x}\frac{\partial F}{\partial u}\,\mathrm{d}u + \left(\frac{\partial v}{\partial x}\frac{\partial F}{\partial y} + \frac{\partial v}{\partial y}\right)\mathrm{d}y$$

$$= \frac{\partial v}{\partial x}\frac{\partial F}{\partial u}\,\mathrm{d}u + \left(\frac{\dfrac{\partial u}{\partial x}\dfrac{\partial v}{\partial y} - \dfrac{\partial v}{\partial x}\dfrac{\partial u}{\partial y}}{\dfrac{\partial u}{\partial x}}\right)\mathrm{d}y \qquad \text{using the above}$$

expression for $\dfrac{\partial F}{\partial y}$. The numerator of the second term is precisely $\dfrac{\partial(u, v)}{\partial(x, y)}$ which is zero. Hence $\mathrm{d}v = \dfrac{\partial v}{\partial x}\dfrac{\partial F}{\partial u}\,\mathrm{d}u$. However we have, since v is a function of u and y, $\mathrm{d}v = \dfrac{\partial v}{\partial u}\,\mathrm{d}u + \dfrac{\partial v}{\partial y}\,\mathrm{d}y$. Comparing these two differentials, gives $\dfrac{\partial v}{\partial y} = 0$. This means that v is solely a function of u, that is $v = g(u)$ or $v - g(u) = 0$ which is $\phi(u, v) = 0$ as required. There is no doubt that this is a correct proof, but care needs to be taken in the interpretation of the partial derivatives. For example, $\dfrac{\partial v}{\partial y} = 0$ only if u is being held constant. It is not true if x is being held constant, otherwise we have the nonsense that any function $v = g(x, y)$ can be deemed independent of y!

This result is useful because it gives a good test. Given $u = f(x, y)$ and $v = g(x, y)$, find $\dfrac{\partial(u, v)}{\partial(x, y)}$, and if this is zero, there is a functional relationship between u and v.

Example 2.9 If $u = \sin^{-1} x + \sin^{-1} y$ and $v = x\sqrt{1 - y^2} + y\sqrt{1 - x^2}$ determine whether there is a functional relationship between v and u.

Solution For this kind of problem, the best way to proceed is to determine whether the Jacobian $\dfrac{\partial(u, v)}{\partial(x, y)}$ is zero. $\dfrac{\partial u}{\partial x} = \dfrac{1}{\sqrt{1 - x^2}}$, $\dfrac{\partial u}{\partial y} = \dfrac{1}{\sqrt{1 - y^2}}$, $\dfrac{\partial v}{\partial x} = \sqrt{1 - y^2} - \dfrac{xy}{\sqrt{1 - x^2}}$,

$\dfrac{\partial v}{\partial y} = \sqrt{1 - x^2} - \dfrac{xy}{\sqrt{1 - y^2}}$. Hence, $\dfrac{\partial(u, v)}{\partial(x, y)} = \dfrac{\partial u}{\partial x}\dfrac{\partial v}{\partial y} - \dfrac{\partial u}{\partial y}\dfrac{\partial v}{\partial x}$

$$= \frac{1}{\sqrt{1 - x^2}}\left(\sqrt{1 - x^2} - \frac{xy}{\sqrt{1 - y^2}}\right) - \frac{1}{\sqrt{1 - y^2}}\left(\sqrt{1 - y^2} - \frac{xy}{\sqrt{1 - x^2}}\right)$$

$$= 1 - \frac{xy}{\sqrt{(1 - x^2)(1 - y^2)}} - 1 + \frac{xy}{\sqrt{(1 - x^2)(1 - y^2)}}$$

$$= 0$$

Hence by Example 2.4, there must be a relationship between v and u. In this particular case, it can easily be found. Simply put $\theta = \sin^{-1} x$ and $\phi = \sin^{-1} y$ so that $u = \theta + \phi$, and $v = \sin\theta\cos\phi + \cos\theta\sin\phi = \sin(\theta + \phi) = \sin u$. So $v = \sin u$ and the relationship is confirmed.

Example 2.10 The volume V of a right circular cone, base radius a and height h is given by the expression

$$V = \frac{1}{3}\pi a^2 h$$

In the manufacture of the cone, $a = 2$ m but with an error of 0.01 m and $h = 3$ m with an error of 0.02 m. Find the value of V and an approximation to its greatest possible error.

Solution In order to solve this problem, the total differential is used so that, since

$$V = V(a,h), \qquad dV = \frac{\partial V}{\partial a}\,da + \frac{\partial V}{\partial h}\,dh$$

However, to be entirely correct, we replace dV by ΔV etc. So

$$\Delta V \approx \frac{\partial V}{\partial a}\Delta a + \frac{\partial V}{\partial h}\Delta h$$

(Differentials are zero (infinitesimal), ΔV etc. are small but finite.)
Now $\dfrac{\partial V}{\partial a} = \dfrac{2}{3}\pi ah$ and $\dfrac{\partial V}{\partial h} = \dfrac{1}{3}\pi a^2$ and we are given maximum values for Δa and Δh as 0.01 and 0.02 respectively. This then enables us to calculate an approximation to the largest possible error for ΔV:

$$\Delta V \approx \frac{2}{3}\pi \times 2 \times 3 \times 0.01 + \frac{1}{3}\pi \times 2^2 \times 0.02$$
$$= 0.21 \text{ m}^3$$

The value of V therefore will vary by this amount about the central value given by

$$V = \frac{1}{3}\pi a^2 h = \frac{1}{3}\pi \times 2^2 \times 3 = 4\pi = 12.57$$

Example 2.11 The volume flow of water along a pipe, V, is given by

$$V = \frac{\pi p a^4}{8\mu l}$$

(p = pressure, a = radius of pipe, μ = dynamic viscosity, l = length of pipe.)
If a and l both increase by 5 per cent, p decreases by 10 per cent and μ decreases by 30 per cent, find the percentage change in V.

Solution This time, $V = V(p, a, \mu, l)$ so the differential $\mathrm{d}V$ is given by

$$\mathrm{d}V = \frac{\partial V}{\partial p}\mathrm{d}p + \frac{\partial V}{\partial a}\mathrm{d}a + \frac{\partial V}{\partial \mu}\mathrm{d}\mu + \frac{\partial V}{\partial l}\mathrm{d}l$$

$$\frac{\partial V}{\partial p} = \frac{\pi a^4}{8\mu l},\ \frac{\partial V}{\partial a} = \frac{\pi p a^3}{2\mu l},\ \frac{\partial V}{\partial \mu} = -\frac{\pi p a^4}{8\mu^2 l},\ \frac{\partial V}{\partial l} = -\frac{\pi p a^4}{8\mu l^2}$$

Dividing by V gives

$$\frac{\mathrm{d}V}{V} = \frac{1}{V}\frac{\partial V}{\partial p}\mathrm{d}p + \frac{1}{V}\frac{\partial V}{\partial a}\mathrm{d}a + \frac{1}{V}\frac{\partial V}{\partial \mu}\mathrm{d}\mu + \frac{1}{V}\frac{\partial V}{\partial l}\mathrm{d}l$$

$$\frac{\mathrm{d}V}{V} = \frac{\mathrm{d}p}{p} + 4\frac{\mathrm{d}a}{a} - \frac{\mathrm{d}\mu}{\mu} - \frac{\mathrm{d}l}{l}$$

using the above formulae for the partial derivatives.

Interestingly, if we take the formula $V = \dfrac{\pi p a^4}{8\mu l}$ and take logarithms we obtain

$$\ln V = \ln\left(\frac{\pi}{8}\right) + \ln p + 4\ln a - \ln \mu - \ln l$$

Differentiating and since $\dfrac{\mathrm{d}}{\mathrm{d}V}(\ln V) = \dfrac{1}{V}$ so $\mathrm{d}(\ln V) = \dfrac{\mathrm{d}V}{V}$, $\mathrm{d}(\ln p) = \dfrac{\mathrm{d}p}{p}$ etc. re-derives the formula for $\dfrac{\mathrm{d}V}{V}$. Sometimes the phrase 'logarithmic differentiation' is used for this process.

Using the data in the question we have

$$\begin{aligned}\frac{\Delta V}{V} &\approx \frac{\Delta p}{p} + 4\frac{\Delta a}{a} - \frac{\Delta \mu}{\mu} - \frac{\Delta l}{l}\\ &= -0.1 + 4 \times 0.05 + 0.3 - 0.05\\ &= 0.35\end{aligned}$$

Hence V experiences a 35 per cent increase.

This is, of course, approximate since the exact differentials have been replaced by small but non-zero quantities. Precise discussion of accuracy is out of place here.

Example 2.12 If $F(\alpha) = \int_{a(\alpha)}^{b(\alpha)} f(x, \alpha)\mathrm{d}x$ show that

$$\frac{\mathrm{d}F}{\mathrm{d}\alpha} = \int_{a(\alpha)}^{b(\alpha)} \frac{\partial f}{\partial \alpha}\mathrm{d}x + f(b, \alpha)\frac{\mathrm{d}b}{\mathrm{d}\alpha} - f(a, \alpha)\frac{\mathrm{d}a}{\mathrm{d}\alpha}$$

where a, b, f and F are all differentiable with respect to α.

Solution We give two proofs of this useful result (often called Leibniz's Rule).

Proof 1

This proof uses the definition and Mean Value Theorem (see Example 1.8) but applied to integrals.

By definition of differentiation of a single variable

$$\frac{\mathrm{d}F}{\mathrm{d}\alpha} = \lim_{h\to 0} \frac{F(\alpha + h) - F(\alpha)}{h}$$

where the numerator is the expression

$$\int_{a(\alpha+h)}^{b(\alpha+h)} f(x, \alpha + h)\mathrm{d}x - \int_{a(\alpha)}^{b(\alpha)} f(x, \alpha)\mathrm{d}x$$

We now manipulate the first term to generate differences that can be utilised by the Mean Value Theorem. First of all, note that

$$\int_{a(\alpha)}^{b(\alpha+h)} = \int_{a(\alpha)}^{a(\alpha+h)} + \int_{a(\alpha+h)}^{b(\alpha+h)}$$

and $\displaystyle\int_{a(\alpha)}^{b(\alpha+h)} = \int_{a(\alpha)}^{b(\alpha)} + \int_{b(\alpha)}^{b(\alpha+h)}$

so $\displaystyle\int_{a(\alpha+h)}^{b(\alpha+h)} = \int_{a(\alpha+h)}^{a(\alpha)} + \int_{a(\alpha)}^{b(\alpha+h)} \quad \left(\text{using } \int_{a(\alpha)}^{a(\alpha+h)} = -\int_{a(\alpha+h)}^{a(\alpha)}\right)$

$$= \int_{a(\alpha+h)}^{a(\alpha)} + \int_{a(\alpha)}^{b(\alpha)} + \int_{b(\alpha)}^{b(\alpha+h)}$$

Hence $\displaystyle\int_{a(\alpha+h)}^{b(\alpha+h)} f(x, \alpha + h)\mathrm{d}x - \int_{a(\alpha)}^{b(\alpha)} f(x, \alpha)\mathrm{d}x$

$$= \int_{a(\alpha)}^{b(\alpha)} (f(x, \alpha + h) - f(x, \alpha))\mathrm{d}x + \int_{b(\alpha)}^{b(\alpha+h)} f(x, \alpha + h)\mathrm{d}x - \int_{a(\alpha)}^{a(\alpha+h)} f(x, \alpha + h)\mathrm{d}x$$

Now $\displaystyle\int_{b(\alpha)}^{b(\alpha+h)} f(x, \alpha + h)\mathrm{d}x = f(\xi_1, \alpha + h)[b(\alpha + h) - b(\alpha)]$

and $\displaystyle\int_{a(\alpha)}^{a(\alpha+h)} f(x, \alpha + h)\mathrm{d}x = f(\xi_2, \alpha + h)[a(\alpha + h) - a(\alpha)]$

where $b(\alpha) \le \xi_1 \le b(\alpha + h)$, $a(\alpha) \le \xi_2 \le a(\alpha + h)$
using the first Mean Value Theorem.

Thus $\displaystyle\frac{\mathrm{d}F}{\mathrm{d}\alpha} = \lim_{h\to 0} \left\{\frac{F(\alpha + h) - F(\alpha)}{h}\right\}$

$$= \lim_{h\to 0} \left[\begin{aligned} &\int_{a(\alpha)}^{b(\alpha)} \frac{1}{h}(f(x, \alpha + h) - f(x, \alpha))\mathrm{d}x + \frac{f(\xi_1, \alpha + h)}{h}(b(\alpha + h) - b(\alpha)) \\ &- \frac{f(\xi_2, \alpha + h)}{h}(a(\alpha + h) - a(\alpha)) \end{aligned}\right]$$

$$= \int_{a(\alpha)}^{b(\alpha)} \frac{\partial f}{\partial \alpha}\mathrm{d}x + f(b, \alpha)\frac{\mathrm{d}b}{\mathrm{d}\alpha} - f(a, \alpha)\frac{\mathrm{d}a}{\mathrm{d}\alpha}$$

since ξ_1 and ξ_2 tend to b and a respectively as $h \to 0$. This establishes the result.

Proof 2

This proof is shorter, but perhaps less rigorous.

Since $F = \int_{a(\alpha)}^{b(\alpha)} f(x, \alpha)\mathrm{d}x$, F can be thought of as a function of not only α, but b and a too. Hence we can write

$$\frac{\mathrm{d}F}{\mathrm{d}\alpha} = \frac{\partial F}{\partial \alpha} + \frac{\partial F}{\partial b}\frac{\mathrm{d}b}{\mathrm{d}\alpha} + \frac{\partial F}{\partial a}\frac{\mathrm{d}a}{\mathrm{d}\alpha}$$

because a and b are functions of α. (This is the chain rule from $F(\alpha, b, a)$ to $F(\alpha)$.)

Now the fundamental rule of calculus states that

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_a^x f(t)\mathrm{d}t = f(x)$$

Hence $\displaystyle\frac{\partial F}{\partial b} = \frac{\partial}{\partial b}\int_a^b f(x, \alpha)\mathrm{d}x = f(b, \alpha)$

and $\dfrac{\partial F}{\partial a} = \dfrac{\partial}{\partial a}\displaystyle\int_a^b f(x, \alpha)\mathrm{d}x = -f(a, \alpha)$

Interpreting $\dfrac{\partial F}{\partial \alpha}$ as 'differentiating F holding a and b constant' gives

$$\frac{\partial F}{\partial \alpha} = \int_a^b \frac{\partial f}{\partial \alpha}\,\mathrm{d}x$$

whence $$\frac{\mathrm{d}F}{\mathrm{d}\alpha} = \int_{a(\alpha)}^{b(\alpha)} \frac{\partial f}{\partial \alpha}\,\mathrm{d}x + f(b, \alpha)\frac{\mathrm{d}b}{\mathrm{d}\alpha} - f(a, \alpha)\frac{\mathrm{d}a}{\mathrm{d}\alpha}$$

Example 2.13 Evaluate the integral

$$\int_{-\pi}^{\pi} \frac{\mathrm{d}x}{(7 + \sin x)^2}$$

given $\displaystyle\int_{-\pi}^{\pi} \frac{\mathrm{d}x}{\alpha + \sin x} = \frac{2\pi}{\sqrt{\alpha^2 - 1}}$ $\qquad (\alpha > 1)$

Solution Using Leibniz's Rule; if

$$\int_{-\pi}^{\pi} \frac{\mathrm{d}x}{\alpha + \sin x} = \frac{2\pi}{\sqrt{\alpha^2 - 1}}$$

differentiating both sides with respect to α gives

$$-\int_{-\pi}^{\pi} \frac{\mathrm{d}x}{(\alpha + \sin x)^2} = -2\pi \cdot 2\alpha \cdot \tfrac{1}{2}(\alpha^2 - 1)^{-\frac{3}{2}}$$

so $$\int_{-\pi}^{\pi} \frac{\mathrm{d}x}{(\alpha + \sin x)^2} = 2\pi\alpha(\alpha^2 - 1)^{-\frac{3}{2}}$$

Let $\alpha = 7$ to obtain $\displaystyle\int_{-\pi}^{\pi} \frac{\mathrm{d}x}{(7 + \sin x)^2} = 14\pi(48)^{-\frac{3}{2}} = \frac{14\pi}{48\sqrt{48}} = \frac{7\pi}{96\sqrt{3}}$

Example 2.14 Evaluate the integral

$$\int_0^1 x^p \ln x\mathrm{d}x$$

given $\displaystyle\int_0^1 x^p\mathrm{d}x = \frac{1}{p + 1}$

and deduce the value of $\displaystyle\int_0^1 x^p (\ln x)^n\mathrm{d}x$.

Solution Differentiating with respect to p, using the standard form

$$\frac{\mathrm{d}}{\mathrm{d}p}(x^p) = x^p\ln x, \text{ gives}$$

$$\int_0^1 x^p \ln x\mathrm{d}x = \frac{\mathrm{d}}{\mathrm{d}p}\left(\frac{1}{p + 1}\right) = -\frac{1}{(p + 1)^2}$$

Differentiating n times, gives

$$\int_0^1 x^p(\ln x)^n\mathrm{d}x = (-1)^n \frac{n!}{(p + 1)^{n+1}}$$

Example 2.15 If F is a homogeneous function of degree m, and $F = F(x_1, x_2, \ldots, x_n)$, show that $\sum_{i=1}^{n} x_i \dfrac{\partial f}{\partial x_i} = mF$.

Hence deduce the value of $\sum_{i=1}^{n} \dot{q}_i \dfrac{\partial T}{\partial \dot{q}_i}$, where T is the kinetic energy of n independent particles, the ith of which has speed $\dot{q}_i$.

Solution Let $U = F(tx_1, tx_2, \ldots, tx_n)$. By definition of homogeneous function, U is also given by $U = t^m F(x_1, x_2, \ldots, x_n)$.
Now suppose $(x_1, x_2, \ldots, x_n)$ does not depend on t so that

$$\frac{\mathrm{d}U}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}(kt^m) = mkt^{m-1}$$

Writing $y_i = tx_i, \quad i = 1, 2, \ldots, n$

$$U = F(y_1, y_2, \ldots, y_n) = F(t)$$

so that the chain rule gives

$$\frac{\mathrm{d}U}{\mathrm{d}t} = \frac{\partial F}{\partial y_1}\frac{\mathrm{d}y_1}{\mathrm{d}t} + \frac{\partial F}{\partial y_2}\frac{\mathrm{d}y_2}{\mathrm{d}t} + \ldots + \frac{\partial F}{\partial y_n}\frac{\mathrm{d}y_n}{\mathrm{d}t}$$

hence

$$\frac{\mathrm{d}U}{\mathrm{d}t} = x_1\frac{\partial F}{\partial y_1} + x_2\frac{\partial F}{\partial y_2} + \ldots + x_n\frac{\partial F}{\partial y_n}$$

Also $\dfrac{\mathrm{d}U}{\mathrm{d}t} = mkt^{m-1}$
Multiply both results by t to give

$$mF = y_1\frac{\partial F}{\partial y_1} + y_2\frac{\partial F}{\partial y_2} + \ldots + y_n\frac{\partial F}{\partial y_n} \tag{I}$$

which proves the result.
T, the kinetic energy, takes the form of a quadratic function

$$T = \tfrac{1}{2}\sum_{i=1}^{n} m_i \dot{q}_i^2$$

Hence substituting $m = 2$, $y_i = \dot{q}_i$ in result (I) gives

$$2T = \dot{q}_1\frac{\partial T}{\partial \dot{q}_1} + \ldots + \dot{q}_n\frac{\partial T}{\partial \dot{q}_n}$$

so

$$\sum_{i=1}^{n} \dot{q}_i\frac{\partial T}{\partial \dot{q}_i} = 2T$$

(This result is used in Lagrangian mechanics – see the author's *Mechanics Work Out* in this series.)

Example 2.16 Show that, if f and g are twice differentiable functions of x and t, then the differential equation

$$\frac{\partial^2\phi}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2}$$

where c is a constant, is satisfied by the function $\phi(x, t) = f(x + ct) + g(x - ct)$.

Solution Let $\xi = x + ct$ and $\eta = x - ct$, since these combinations appear in the expression for $\phi(x, t)$. Using the chain rule:

$$\frac{\partial}{\partial x} = \frac{\partial \xi}{\partial x}\frac{\partial}{\partial \xi} + \frac{\partial \eta}{\partial x}\frac{\partial}{\partial \eta}$$

so
$$\frac{\partial}{\partial x} = \frac{\partial}{\partial \xi} + \frac{\partial}{\partial \eta}$$

and
$$\frac{\partial}{\partial t} = \frac{\partial \xi}{\partial t}\frac{\partial}{\partial \xi} + \frac{\partial \eta}{\partial t}\frac{\partial}{\partial \eta}$$

so
$$\frac{\partial}{\partial t} = c\left(\frac{\partial}{\partial \xi} - \frac{\partial}{\partial \eta}\right)$$

The application of these operator equations gives

$$\frac{\partial^2 \phi}{\partial x^2} = \left(\frac{\partial}{\partial \xi} + \frac{\partial}{\partial \eta}\right)\left(\frac{\partial \phi}{\partial \xi} + \frac{\partial \phi}{\partial \eta}\right)$$
$$= \frac{\partial^2 \phi}{\partial \xi^2} + 2\frac{\partial^2 \phi}{\partial \xi \partial \eta} + \frac{\partial^2 \phi}{\partial \eta^2}$$

and
$$\frac{\partial^2 \phi}{\partial t^2} = c^2\left(\frac{\partial}{\partial \xi} - \frac{\partial}{\partial \eta}\right)\left(\frac{\partial \phi}{\partial \xi} - \frac{\partial \phi}{\partial \eta}\right)$$
$$= c^2\left(\frac{\partial^2 \phi}{\partial \xi^2} - 2\frac{\partial^2 \phi}{\partial \xi \partial \eta} + \frac{\partial^2 \phi}{\partial \eta^2}\right)$$

So
$$\frac{\partial^2 \phi}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 \phi}{\partial t^2} = 4\frac{\partial \phi^2}{\partial \xi \partial \eta} = 0$$

Integrating $\dfrac{\partial \phi^2}{\partial \xi \partial \eta} = 0$ with respect to η gives

$$\frac{\partial \phi}{\partial \xi} = \alpha(\xi)$$

where $\alpha(\xi)$ is an arbitrary function of ξ. Integrating this with respect to ξ gives

$$\phi = \int \alpha(\xi)\mathrm{d}\xi + g(\eta)$$

Or, letting $f(\xi) = \int \alpha(\xi)\mathrm{d}\xi$, gives

$$\phi = f(\xi) + g(\eta)$$
$$\phi(x, t) = f(x + ct) + g(x - ct)$$

Thus $\phi(x, t)$ satisfies $\dfrac{\partial^2 \phi}{\partial x^2} = \dfrac{1}{c^2}\dfrac{\partial^2 \phi}{\partial t^2}$.

Example 2.17 Given that $\int_0^\infty \mathrm{e}^{-ax}\sin bx\mathrm{d}x = \dfrac{b}{a^2 + b^2}$ where a and b are constants, $a > 0$, evaluate the integrals $\int_0^\infty x\mathrm{e}^{-ax}\sin bx\mathrm{d}x$, and $\int_0^\infty \dfrac{\mathrm{e}^{-ax}\sin bx}{x}\mathrm{d}x$ by using Leibniz's Rule and its extension. Hence deduce the value of $\int_0^\infty \dfrac{\sin x}{x}\mathrm{d}x$.

Solution Starting with the result $\int_0^\infty \mathrm{e}^{-ax}\sin bx\mathrm{d}x = \dfrac{b}{a^2 + b^2}$, we differentiate both sides with respect to a to give $\int_0^\infty -x\mathrm{e}^{-ax}\sin bx\mathrm{d}x = \dfrac{\mathrm{d}}{\mathrm{d}a}\left(\dfrac{b}{a^2 + b^2}\right) = \dfrac{-2ab}{(a^2 + b^2)^2}$. Whence $\int_0^\infty x\mathrm{e}^{-ax}\sin bx\mathrm{d}x = \dfrac{2ab}{(a^2 + b^2)^2}$. For the second integral, we use *integration* under the integral sign, an extension of Leibniz's Rule.

Again starting with $\int_0^\infty e^{-ax}\sin bx\,dx = \dfrac{b}{a^2+b^2}$, integrate both sides with respect to a between the limits 0 and a to obtain $\int_0^a\int_0^\infty e^{-ax}\sin bx\,dx\,da = \int_0^a \dfrac{b}{a^2+b^2}\,da = \tan^{-1}\left(\dfrac{a}{b}\right)$ using a standard integral. The two integrals on the left-hand side can be interchanged since they both converge to give $\int_0^\infty\left[-\dfrac{1}{x}e^{-ax}\right]_{a=0}^{a=a}\sin bx\,dx = \int_0^\infty\left(\dfrac{1-e^{-ax}}{x}\right)\sin bx\,dx = \tan^{-1}\left(\dfrac{a}{b}\right)$. This is true for all values of b and for all positive values of a. We can therefore let $a\to\infty$ to deduce the result

$$\int_0^\infty\left(\frac{1}{x}\right)\sin bx\,dx = \lim_{a\to\infty}\left(\tan^{-1}\left(\frac{a}{b}\right)\right) = \tan^{-1}(\infty) = \frac{\pi}{2}$$

This result is independent of the value of b which will come as no surprise if you substitute $u = bx$, hence $\int_0^\infty \dfrac{\sin x}{x}\,dx = \dfrac{\pi}{2}$, a result that is frequently used in signal processing and other branches of electronics. Going back to the previous result $\int_0^\infty\left(\dfrac{1-e^{-ax}}{x}\right)\sin bx\,dx = \tan^{-1}\left(\dfrac{a}{b}\right)$ we can therefore deduce that $\int_0^\infty\left(\dfrac{1-e^{-ax}}{x}\right)\sin bx\,dx = \int_0^\infty\left(\dfrac{1}{x}\right)\sin bx\,dx - \int_0^\infty\dfrac{e^{-ax}}{x}\sin bx\,dx = \tan^{-1}\left(\dfrac{a}{b}\right)$ as all integrals converge. However we have just derived $\int_0^\infty\left(\dfrac{1}{x}\right)\sin bx\,dx = \dfrac{\pi}{2}$ so we have, finally

$$\int_0^\infty\frac{e^{-ax}}{x}\sin bx\,dx = \frac{\pi}{2} - \tan^{-1}\left(\frac{a}{b}\right)$$

To evaluate this integral any other way would be challenging!

2.3 Exercises

2.1. A function of two variables is defined by $f(x, y) = x^4 + 5x^3y + 5xy^3 + y^4$. Find the partial derivatives $f_x, f_y, f_{xx}, f_{yy}, f_{xy}$.

2.2. Find $f_x, f_y, f_z, f_{xx}, f_{yy}, f_{zz}, f_{xy}, f_{yz}, f_{zx}$ if f is a function of three variables defined by $f(x, y, z) = (x + y + z)e^{(x+y-z)}$.

2.3. For the following pairs of functions $u(x, y)$, $v(x, y)$ show that $\dfrac{\partial u}{\partial x} = \dfrac{\partial v}{\partial y}$ and $\dfrac{\partial u}{\partial y} = -\dfrac{\partial v}{\partial x}$ (called the *Cauchy–Riemann equations*).
(a) $u = x^2 - y^2$, $v = 2xy$,
(b) $u = e^x\cos y$, $v = e^x\sin y$,
(c) $u = \frac{1}{2}\ln(x^2 + y^2)$, $v = \tan^{-1}\left(\dfrac{y}{x}\right)$,
(d) $u = \dfrac{x}{x^2+y^2}$, $v = -\dfrac{y}{x^2+y^2}$.

Show further that in each case, $u_{xx} + u_{yy} = 0$ and $v_{xx} + v_{yy} = 0$ (that is, both u and v are *harmonic* functions).

2.4. Show that the function $f(x, y) = Ax^4 + Bx^2y^2 + Cy^4$ satisfies the partial differential equation $x\dfrac{\partial f}{\partial x} + y\dfrac{\partial f}{\partial y} = 4f$.

2.5. In polar co-ordinates, $x = R\cos\theta$, $y = R\sin\theta$. Find $\dfrac{\partial x}{\partial R}\dfrac{\partial y}{\partial \theta} - \dfrac{\partial x}{\partial \theta}\dfrac{\partial y}{\partial R}$, the Jacobian.

2.6. The ideal gas law states that $P = k\dfrac{T}{V}$, where k = constant, T = temperature, P = pressure and V = volume. Show that $V\dfrac{\partial P}{\partial V} = -P$, and $V\dfrac{\partial P}{\partial V} + T\dfrac{\partial P}{\partial T} = 0$.

2.7. Let $z = \dfrac{x+y}{x-y}$ where $x = \cos t$ and $y = \sin t$. Use the chain rule to find $\dfrac{dz}{dt}$, and check your result by direct differentiation.

2.8. Show that if $f(x, y) = \dfrac{xy}{x+y}$ then $x^2f_{xx} + 2xyf_{xy} + y^2f_{yy} = 0$.

2.9. If $\phi = F\left(\dfrac{y}{x}\right)$ prove that $x\dfrac{\partial\phi}{\partial x} + y\dfrac{\partial\phi}{\partial y} = 0$.

2.10. The function ϕ satisfies the partial differential equation

$$c\,\frac{\partial^2\phi}{\partial x^2} = \frac{\partial^2\phi}{\partial x\partial t}.$$

Show that the general solution takes the form $\phi = f(x + ct) + g(t)$ where f and g are arbitrary functions.

2.11. Determine $\dfrac{\partial r}{\partial x_i}$, $i = 1, 2, \ldots, n$ where $r^2 = x_1^2 + x_2^2 + \ldots + x_n^2$. Hence show that $\sum_{i=1}^{n} x_i\dfrac{\partial r}{\partial x_i} = r$.

2.12. Given that $w^2 + w\sin xyz = 1$, find $\dfrac{\partial w}{\partial x}$, $\dfrac{\partial w}{\partial y}$ and $\dfrac{\partial w}{\partial z}$.

2.13. If $x = e^u\sec v$ and $y = e^u\tan v$ show that

$$\left(\frac{\partial f}{\partial x}\right)^2 - \left(\frac{\partial f}{\partial y}\right)^2 = e^{-2u}\left\{\left(\frac{\partial f}{\partial u}\right)^2 - \cos^2 v\left(\frac{\partial f}{\partial v}\right)^2\right\}$$

where $f(x, y)$ is a differentiable function of two variables.

2.14. By writing $u = x + y$, $v = xy$ transform the equation

$$(x - y)(x^2\phi_{xx} - 2xy\phi_{xy} + y^2\phi_{yy}) = 2xy(\phi_x - \phi_y)$$

into one involving derivatives of ϕ with respect to u and v. Hence solve it.

2.15. If $u = \tan^{-1}x + \tan^{-1}y$ and $v = \dfrac{x + y}{1 - xy}$ determine whether there is a functional relationship between u and v by finding the Jacobian $\dfrac{\partial(u, v)}{\partial(x, y)}$.

2.16. The volume of a frustum of a cone shown in Figure 2.1 is given by

$$V = \tfrac{1}{3}\pi h(b^2 + ba + a^2)$$

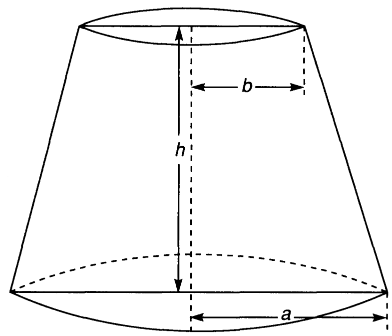


Figure 2.1 The frustum of a cone. (Note that $a > b$, but this is not essential.)

Find the rates of change of V with respect to b, a and h. What is the error in V if b, a, h can have errors of 1 per cent, 2 per cent and 3 per cent respectively, and $a = 0.5b$.

2.17. The area of a triangle ABC with sides a, b, c is given by $\frac{1}{2}ab\sin C$. The notation is standard, but see Figure 2.2 if schooldays are too distant a memory.

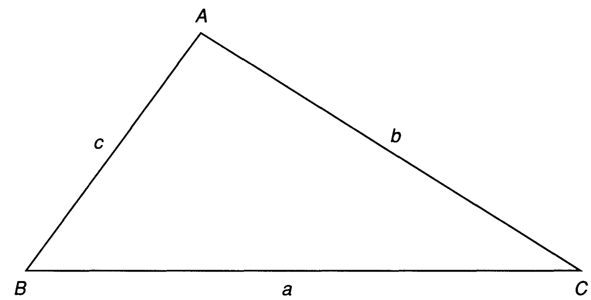


Figure 2.2 The triangle ABC.

The errors in a and b are both 2 per cent. The angle C can be in error by 5 degrees. Calculate the maximum error in the area of the triangle if $C = 45$ degrees.

2.18. Use the result $\int_0^\infty e^{-ax}\cos bx\,dx = \dfrac{a}{a^2 + b^2}$ $(a > 0)$, to find the values of the integrals $\int_0^\infty xe^{-ax}\cos bx\,dx$ and $\int_0^\infty xe^{-ax}\sin bx\,dx$.

2.19. Show that $\int_0^\infty e^{-st}\,dt = \dfrac{1}{s}$ and hence deduce the value of $\int_0^\infty t^n e^{-st}\,dt$ where n is an integer.

Topic Guide

Taylor's Theorem in Two Variables
Criteria for Max. and Min.
Extension to Many Variables
Constraints
Lagrange Multipliers

3 Maxima and Minima

3.1 Fact Sheet

A function of two variables $f(x, y)$ when written in the form $z = f(x, y)$ represents a surface in three-dimensional space; it is this that provides the principal application of the theory that follows. Taylor's Theorem in two variables takes the form:

$$f(a+h, b+k) = f(a, b) + \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)f(a, b) + \frac{1}{2!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^2 f(a, b) + \ldots$$
$$+\frac{1}{n!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^n f(a, b) + R_n$$

where $\frac{1}{r!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^r f(a, b)$ is interpreted as the operator $\frac{1}{r!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^r$ acting on the function $f(x, y)$ then x placed equal to a and y placed equal to b ($r = 1, 2, \ldots, n$). R_n is the remainder term.

Taylor's Theorem in n variables has the general form:

$$F = f + \sum_{i=1}^{n} h_i f_i + \frac{1}{2!}\sum_{i=1}^{n}\sum_{j=1}^{n} h_i h_j f_{ij} + \ldots$$

where $F = f(a_1 + h_1, a_2 + h_2, \ldots, a_n + h_n)$, the h_i ($i = 1, 2, \ldots, n$) are small, $f_i = \frac{\partial f}{\partial x_i}$, $f_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}$, etc., and f and all its derivatives on the right-hand side are evaluated at $x_i = a_i$. The first (single) summation is a general linear expression since all of the f_i's are simply numbers, the second (double) summation is, similarly, a general quadratic form and so on. Each 'term' is a higher-order homogeneous function. Little use of Taylor's Theorem in more than two dimensions will be made in this text.

Finding maxima and minima in two variables proceeds as follows:

Using the suffix derivative notation $\frac{\partial f}{\partial x} = f_x$, $\frac{\partial^2 f}{\partial x \partial y} = f_{xy}$ etc., given a function of two variables, $f(x, y)$ say, then a necessary condition for the point (a, b) to be an extremum (that is a maximum, minimum, or some other shape where the tangent plane to the surface $z = f(x, y)$ is parallel to the x–y plane) is $f_x = 0, f_y = 0$ at $(x, y) = (a, b)$ [commonly written $f_x(a, b) = 0, f_y(a, b) = 0$]. Further, if $f_{xx} < 0$ and $f_{xx} f_{yy} > f_{xy}^2$ ($x = a, y = b$), then (a, b) is a local maximum; if $f_{xx} > 0$ and again $f_{xx} f_{yy} > f_{xy}^2$ ($x = a, y = b$), then (a, b) is a local minimum; if at $(x, y) = (a, b)$, $f_{xx} f_{yy} < f_{xy}^2$, then (a, b) is termed a saddle point. Other cases need to be investigated from first principles (see Example 3.6).

For a function of n variables, a necessary condition for the point $(x_1, x_2, \ldots, x_n) = (a_1, a_2, \ldots, a_n)$ to be an extremum is that $\frac{\partial f}{\partial x_i} = 0$ at $x_i = a_1$, $i = 1, 2, \ldots, n$. The generalisation of the character of the extremum can also be made as follows: Using Taylor's Theorem for n variables

$$F = f + \sum_{i=1}^{n} h_i f_i + \frac{1}{2!} \sum_{i=1}^{n} \sum_{j=1}^{n} h_i h_j f_{ij} = \ldots$$

the first summation is identically zero since $\frac{\partial f}{\partial x_i} = f_i = 0$ for all i, a necessary condition for an extremum. If the quadratic form $\sum_{i=1}^{n}\sum_{j=1}^{n} h_i h_j f_{ij}$ is positive definite, then the extremum is a minimum. On the other hand, if the quadratic form is negative definite, then the extremum is a maximum. The condition we require is that the quadratic form $\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j$ is positive definite if and only if

$$a_{11} > 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \quad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} > 0, \ldots$$ etc., and it is negative definite if

$$a_{11} < 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} < 0, \quad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} < 0, \ldots$$ etc. (For those not too familiar with with determinants, a short résumé is given in Chapter 4.)

If a function of n variables $f(x_1, x_2, \ldots, x_n)$ has an extremum but only if m of the variables are subject to the *constraints*

$$\begin{aligned} g_1(x_1, x_2, \ldots, x_n) &= 0 \\ g_2(x_1, x_2, \ldots, x_n) &= 0 \\ g_3(x_1, x_2, \ldots, x_n) &= 0 \\ &\vdots \\ &\vdots \\ g_m(x_1, x_2, \ldots, x_n) &= 0 \end{aligned}$$

then we form the combined function

$$\phi = f + \sum_{r=1}^{m} \lambda_r g_r$$

The extrema of the function ϕ, treated as function of the $n + m$ variables $(x_1, x_2, \ldots, x_n, \lambda_1, \lambda_2, \ldots, \lambda_m)$ are those of f subject to the constraints $g_r = 0$, $r = 1, 2, \ldots, m$. The λ_r's are called *undetermined multipliers* (or *Lagrange multipliers*).

3.2 Worked Examples

Example 3.1 Consider the function $f(x, y) = \ln(1 + xy)$. Find the Taylor Series expansion to quadratic order for $f(x, y)$ about the following points: (a) (0, 0), (b) (0, 1), (c) (0, -1) and deduce the expansions about the points (1, 0) and (-1, 0). Is there an expansion about the point (1, –1)?

Solution When finding Taylor Series expansions, especially several about different points as required here, it is wise to find all first-order and second-order partial derivatives. These are:

$$f_x = \frac{y}{1+xy}, \quad f_y = \frac{x}{1+xy}, \quad f_{xx} = -\frac{y^2}{(1+xy)^2}, \quad f_{yy} = -\frac{x^2}{(1+xy)^2}$$

and

$$f_{xy} = \frac{1}{(1+xy)^2}.$$

The expansion about the origin (0, 0) required for part (a) follows immediately from the single-variable Maclaurin Series expansion for $\ln(1+t)$, namely.

$$\ln(1+t) = t - \frac{t^2}{2} + \ldots$$

whence

$$\ln(1+xy) = xy - \frac{(xy)^2}{2} + \ldots$$

In non-operator form up to and including quadratic terms, Taylor's Theorem in two variables is

$$f(x, y) \approx f(a, b) + (x-a)f_x + (y-b)f_y + \tfrac{1}{2}\left((x-a)^2 f_{xx} + 2(x-a)(y-b)f_{xy} + (y-b)^2 f_{yy}\right)$$

where all the partial derivatives are evaluated at the point (a, b).

(b) If $a = 0$, $b = 1$ we get

$$\begin{aligned}\ln(1+xy) &\approx 0 + (x-0)\cdot 1 + (y-1)\cdot 0 + \tfrac{1}{2}\left((x-0)^2(-1) + 2(x-0)(y-1) + (y-1)^2\cdot 0\right)\\ &= x + \tfrac{1}{2}\left(x^2 + 2x(y-1)\right)\end{aligned}$$

(c) If $a = 0$, $b = -1$ we get

$$\begin{aligned}\ln(1+xy) &\approx 0 + (x-0)(-1) + (y-1)\cdot 0 + \tfrac{1}{2}\left((x-0)^2(-1) + 2(x-0)(y-1)\cdot 1 + (y-1)^2\cdot 0\right)\\ &= -x + \tfrac{1}{2}\left(x^2 + 2x(y-1)\right)\end{aligned}$$

Since $f(x, y) = \ln(1+xy) = f(y, x)$, the expansions about (1, 0) and (−1, 0) can be deduced immediately simply by swapping the roles of x and y in the expansions already derived in parts (b) and (c). Thus the expansions are
$\ln(1+xy) \approx y + \frac{1}{2}(y^2 + 2y(x-1))$ about the point (1, 0), and
$\ln(1+xy) \approx -y + \frac{1}{2}(y^2 + 2y(x-1))$ about the point (−1, 0). Remember that there are third- and higher-order terms that have been ignored in all of these expansions. Also, the expressions above can be simplified, but we have chosen to keep them in the form of expansions in $(x-a)$ and $(y-b)$. Finally, the function $f(x, y) = \ln(1+xy)$ has an essential singularity wherever $1 + xy = 0$. Since the values $x = -1$, $y = 1$ satisfy this equation, there can be no power series expansion about this point. A computer-generated picture of the surface $z = \ln(1+xy)$ is given as Figure 3.1 at the top of page 37.

Example 3.2 Find the general Taylor Series expansion about an arbitrary point (u_1, u_2, u_3, u_4) for the function of four variables $f(x_1, x_2, x_3, x_4) = x_1x_2x_3x_4$. Hence deduce the Taylor Series expansion about the arbitrary point $(u_1, u_2, \ldots, u_n)$ for the n-variable function $f(x_1, x_2, \ldots, x_n) = x_1x_2\ldots x_n$.

Solution This question is a little theoretical, but gives some experience in using Taylor's Theorem in many variables. Differentiating f gives the following expressions:
$f_{x_1} = x_2x_3x_4, f_{x_2} = x_1x_3x_4, f_{x_3} = x_1x_2x_4$ and $f_{x_4} = x_1x_2x_3$. The symmetry here is apparent, and it carries forward to the second-order partial derivatives, so for example $f_{x_2x_3} = x_1x_4$ and similarly the third-order derivative $f_{x_2x_3x_4} = x_1$ etc. All fourth-order derivatives are unity and higher-order derivatives are zero. This should not be surprising since the $(n+1)$th-order derivatives of an nth-order polynomial are zero and what we have in this question is a fourth-order polynomial. Once the derivatives are found, we set $x_1 = u_1$, $x_2 = u_2$, $x_3 = u_3$ and $x_4 = u_4$. Also, instead of h and k we use the notation of the fact sheet, namely h_1, h_2, h_3, h_4 to give the Taylor Series

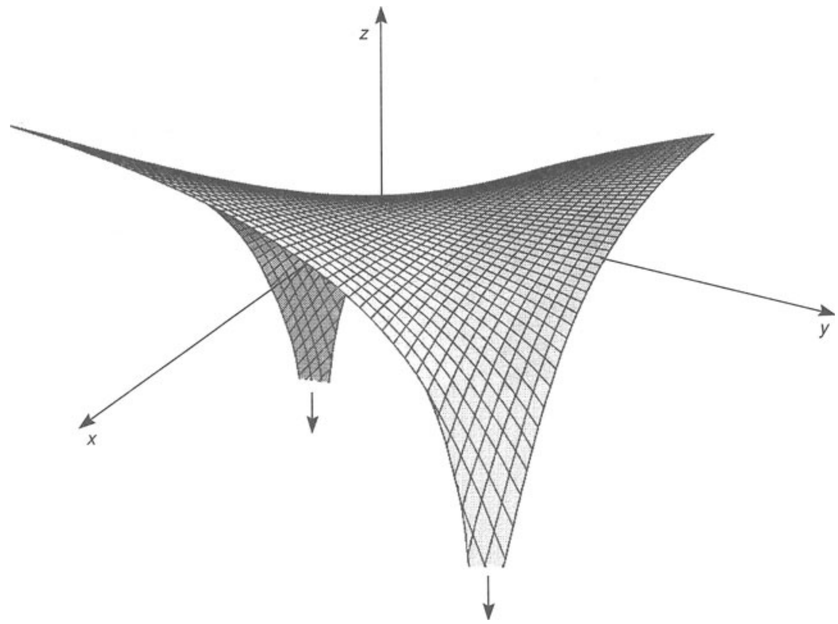


Figure 3.1 The surface $z = \ln(1 + xy)$ (singular at $(1, -1)$ and $(-1, 1)$).

$$\begin{aligned} x_1x_2x_3x_4 &= (u_1 + h_1)(u_2 + h_2)(u_3 + h_3)(u_4 + h_4) \\ &= u_1u_2u_3u_4 + h_1u_2u_3u_4 + h_2u_1u_3u_4 + h_3u_1u_2u_4 + h_4u_1u_2u_3 \\ &\quad + h_1h_2u_3u_4 + \ldots + h_1h_2h_3h_4 \end{aligned}$$

where the factors involving factorials do not appear because of cancellation with binomial type coefficients in the evaluation of $\left(h_1 \frac{\partial}{\partial x_1} + h_2 \frac{\partial}{\partial x_2} + h_3 \frac{\partial}{\partial x_3} + h_4 \frac{\partial}{\partial x_4}\right)^n (x_1x_2x_3x_4)$, n being 1, 2, 3 or 4. Try calculating the third-derivative terms if you need convincing of this (as always, computer algebra is useful here). The Taylor Series in this truncated form is sometimes referred to as a *Taylor polynomial* which is mentioned in Chapter 4 in connection with optimisation. However its principal use is in numerical analysis which is outside the scope of this Work Out (but see *Work Out Numerical Methods* by Peter Turner). (Note, however, that Taylor polynomials are normally *truncated* Taylor Series, not Taylor Series that are naturally finite.) The Taylor Series expansion for $x_1x_2x_3x_4$ about the point (u_1, u_2, u_3, u_4) can thus be found most easily not by performing any calculus, but by multiplying out the four brackets $(u_1 + h_1)(u_2 + h_2)(u_3 + h_3)(u_4 + h_4)$. Similarly, the Taylor Series expansion for the generalisation into n variables is most easily found by expanding $(u_1 + h_1)(u_2 + h_2) \ldots (u_n + h_n)$ algebraically. There is an important point to make here. The Taylor Series expansion of an appropriately differentiable function about a given point is *unique*. Therefore we are at liberty to find it in any convenient manner, assured that once it is found it is the only one there is.

Example 3.3 If $f(x, y)$ is a function of the two variables x and y with continuous second-order derivatives, prove the following: Provided in all cases that $f_x = f_y = 0$ at $x = a$, $y = b$, then
(i) if $f_{xx} > 0$ and $f_{xx}f_{yy} > f^2_{xy}$, (a, b) is a minimum,
(ii) if $f_{xx} < 0$ and $f_{xx}f_{yy} > f^2_{xy}$, (a, b) is a maximum,
(iii) if $f_{xx}f_{yy} < f^2_{xy}$, (a,b) is a saddle point.
Briefly discuss other cases.

Solution The starting point for this problem is Taylor's Theorem in two variables:

$$f(a + h, b + k) = f(a, b) + \left(h \frac{\partial}{\partial x} + k \frac{\partial}{\partial y}\right) f(a, b) + \frac{1}{2!}\left(h \frac{\partial}{\partial x} + k \frac{\partial}{\partial y}\right)^2 f(a, b) + \ldots$$

$$+ \frac{1}{n!}\left(h \frac{\partial}{\partial x} + k \frac{\partial}{\partial y}\right)^n f(a, b) + R_n$$

where the notation is explained in the fact sheet, $(a + h, b + k)$ is a point adjacent to (a, b), in other words, $h, k << 1$. If (a, b) is an extremum the first-derivative terms are zero and, expanding the second-order terms, and rearranging slightly, gives

$$f(a + h, b + k) - f(a, b) = \tfrac{1}{2}\,[h^2 f_{xx} + 2hkf_{xy} + k^2 f_{yy}] + \ldots$$

Since f_x and f_y are zero, there is of course no doubt that the point (a, b) is an extremum. The left-hand side is the (small) difference between the value of $f(x, y)$ at the extremum itself and $f(x, y)$ close by. If we choose h and k so that they enable $f(a + h, b + k)$ to describe a circle around $f(a, b)$ we should be able to decide whether the quantity $f(a + h, b + k) - f(a, b)$ is either always positive ((a, b) is then a minimum) or always negative ((a, b) is then a maximum) or changes sign ((a, b) is then neither a maximum nor a minimum). To this end, let $h = \varepsilon\cos\theta$, $k = \varepsilon\sin\theta$ where θ varies from 0 to 2π and ε is small.

The right-hand side thus becomes

$$\begin{aligned}&\frac{\varepsilon^2}{2}\left[f_{xx}\cos^2\theta + 2f_{xy}\cos\theta\sin\theta + f_{yy}\sin^2\theta\right]\\ &= \frac{\varepsilon^2}{2f_{xx}}\left[f_{xx}^2\cos^2\theta + 2\,f_{xx}f_{xy}\cos\theta\sin\theta + f_{xx}f_{yy}\sin^2\theta\right]\end{aligned}$$

provided $f_{xx} \neq 0$. Completing the square turns this into

$$\frac{\varepsilon^2}{2f_{xx}}\left[(f_{xx}\cos\theta + f_{xy}\sin\theta)^2 + (f_{xx}f_{yy} - f_{xy}^2)\sin^2\theta\right] \qquad \text{(I)}$$

If $f_{xx}f_{yy} > f_{xy}^2$ then expression (I) is positive or negative as f_{xx} is positive or negative. This tells us something about the nature of the extremum, for if (I) is positive for all θ, then the left-hand side $f(a + h, b + k) - f(a, b)$ is also positive for all θ, whence all points that surround (a, b) have values of $f(x, y)$ larger than $f(a, b)$ which implies that the point (a, b) must be a minimum. Similarly, if $f_{xx} < 0$ then (I) is negative for all values of θ and we deduce that the point (a, b) must be a maximum. This establishes (i) and (ii).

If $f_{xx}f_{yy} < f_{xy}^2$, then expression (I) changes sign as θ varies from 0 to 2π (there is no need to worry about the sign of f_{xx} here, as long as f_{xx} is non-zero). For example, (I) is positive if $\theta = 0$ but negative if $\tan\theta = -\dfrac{f_{xx}}{f_{xy}}$. In this case, the extremum is neither a maximum nor a minimum; it is termed a *saddle point*. The name derives from the shape of the saddle familiar to the equestrian but here it encompasses many other shapes. For example, if (I) changed sign many times as θ increased from 0 to 2π (like, say, 8θ), the shape of $f(x, y)$ near such a point would resemble the central part of an old-fashioned jelly mould.

If $f_{xx} = 0$ then

$$f(a + h, b + k) - f(a, b) \approx \tfrac{1}{2}\,\varepsilon^2[2f_{xy}\cos\theta\sin\theta + f_{yy}\sin^2\theta]$$

which, in general, changes in sign and is thus a saddle point. However if both $f_{xx} = 0$ and $f_{xy} = 0$ then the right-hand side (that is, (I)) is the same sign as f_{yy} which leads to there being a maximum at the point (a, b) if $f_{yy} < 0$ and a minimum at the point (a, b) if $f_{yy} > 0$. If $f_{yy} = 0$, then all the ε^2 terms vanish and ε^3 terms which contain the third derivatives of $f(x, y)$ at the point (a, b) need to be considered. Finally, if $f_{yy} = 0$ and the other two second-order partial derivatives of $f(x, y)$ are non-zero, then once again there is a saddle point unless $f_{xy} = 0$, in which case there is a maximum if $f_{xx} < 0$ and a minimum if $f_{xx} > 0$.

Example 3.4 Find and classify all the extreme values of the function

$$f(x, y) = x^2 + 2y^2 - x^2y.$$

Solution With $f(x, y) = x^2 + 2y^2 - x^2y$, differentiating gives

$$\begin{aligned}f_x &= 2x - 2xy\\ f_y &= 4y - x^2\end{aligned}$$

These must simultaneously be zero. We therefore solve the equations

$$
\begin{aligned}
2x - 2xy &= 0 \\
4y - x^2 &= 0
\end{aligned}
$$

The first equation gives $x = 0$ or $y = 1$. If $x = 0$, the second equation gives $y = 0$, whence $(0, 0)$ is an extreme (or stationary) value. If $y = 1$ then the second equation gives $x^2 = 4$ so $x = \pm 2$ giving two further stationary values $(2, 1)$ and $(-2, 1)$. Note the methodical way we enumerate the stationary values of the function $f(x, y)$. As a check, you should always substitute the values back into the simultaneous equations $f_x = 0, f_y = 0$ just to make sure, especially if there are square roots involved. In this example, the three stationary points are indeed $(0, 0)$, $(2, 1)$ and $(-2, 1)$. The second derivatives are as follows:

$$
\begin{aligned}
f_{xx} &= 2 - 2y \\
f_{xy} &= -2x \\
f_{yy} &= 4
\end{aligned}
$$

At $(0, 0)$, $f_{xx} = 2$, $f_{yy} = 4$ and $f_{xy} = 0$ so $f_{xx}f_{yy} > f_{xy}^2$ with $f_{xx} > 0$ which indicates that $(0, 0)$ is a minimum. At $(2, 1)$, $f_{xx} = 0$, $f_{yy} = 4$ and $f_{xy} = -4$. This falls into the final category in the last example, however in practice the criterion involving the second derivatives of $f(x, y)$ can be derived by dividing by f_{yy} instead of f_{xx}, so the condition $f_{xx}f_{yy} < f_{xy}^2$ can still be used to reason that the point $(2, 1)$ is a saddle point in this example. Of course, if either f_{xx} or f_{yy} is zero, then the product must of course be less than the square of f_{xy} and a saddle point is confirmed. For the point $(-2, 1)$, only f_{xy} changes sign and the reasoning remains precisely as in the previous case, hence this point too is a saddle point. Figure 3.2 shows a three-dimensional graph of this surface.

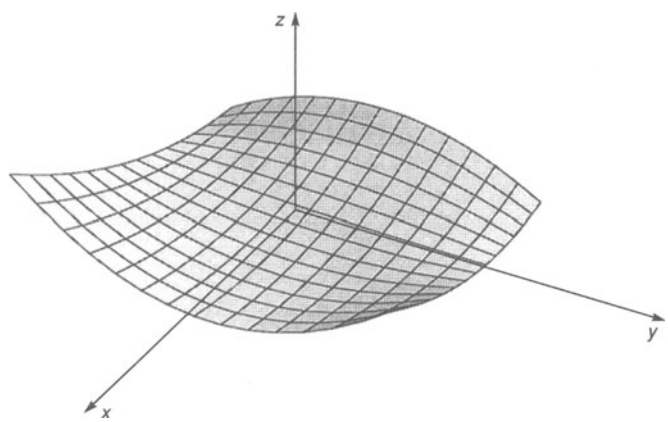


Figure 3.2 The surface $f(x, y) = x^2 + 2y^2 - x^2y$ (drawn for the range $-5 \le x, y \le 5$).

Example 3.5 A function $f(x, y)$ is defined by $f(x, y) = \sin x \cos y$. Find all the extrema of this function and determine whether they are maxima, minima or saddle points.

Solution The first derivatives are given by $f_x = \cos x \cos y$, $f_y = -\sin x \sin y$ hence we solve

$$
\begin{aligned}
\cos x \cos y &= 0 \\
\sin x \sin y &= 0
\end{aligned}
$$

simultaneously. The first of these implies that either $x = (2n + 1)\dfrac{\pi}{2}$ or $y = (2m + 1)\dfrac{\pi}{2}$, and the second implies that either $x = p\pi$ or $y = l\pi$, where p, l, m, and n are integers. Therefore possible extrema are the points $\left((2n + 1)\dfrac{\pi}{2}, p\pi\right)$, $\left(l\pi, (2m + 1)\dfrac{\pi}{2}\right)$. In order to investigate the species, note that the second derivatives are given by

$$
f_{xx} = f_{yy} = -f = -\sin x \cos y, \quad f_{xy} = -\cos x \sin y
$$

The condition that the so-called *discriminant* $f_{xx}f_{yy} - f_{xy}^2$ is positive or negative is therefore particularly easy to check, even at these rather general looking extrema. At the points where $x = (2n + 1)\frac{\pi}{2}$, $y = l\pi$, $f_{xx} = f_{yy} = (-1)^{n+l}$, $f_{xy} = 0$ so the discriminant is positive (= 1) for all values of l and n giving the species of extrema as a maximum if $f_{xx} < 0$, that is $n + l$ is odd, and a minimum if $f_{xx} > 0$, that is $n + l$ is even. At the points where $x = p\pi$, $y = (2m + 1)\frac{\pi}{2}$, $f_{xy}^2 = 1$, $f_{xx} = f_{yy} = 0$ giving the discriminant the value -1 for all values of integers p and m. As pointed out in Example 3.3, although both f_{xx} and f_{yy} are zero and the conditions valid for the derivation of the criterion involving the discriminant have been breached, nevertheless the value -1 for this discriminant still implies that all these extrema are saddle points. If this is troubling the reader, it is easily shown that the Taylor Series taken to second order is $f(x + h, y + k) = f(x, y) - hk(-1)^p(-1)^m$ at these points, which since h and k can be positive or negative means this extremum (technically 'these extrema') is (are) always going to be saddle points. These results equate with common sense; the surface $f(x, y) = \sin x \cos y$ is corrugated like some foam rubber commonly seen in packaging. A drawing is attempted in Figure 3.3.

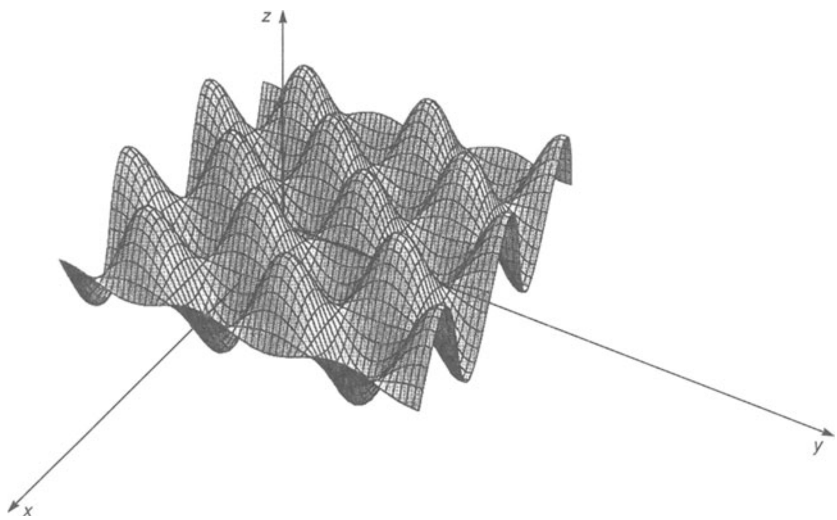


Figure 3.3 The surface $f(x, y) = \sin x \cos y$ (drawn for the range $-8 \le x, y \le 8$).

Example 3.6

Use first principles to investigate the extreme values of the function xye^{-xy}.

Solution

The reason for using first principles for this example is not a perverse whim, but because it is in fact the easiest way to tackle this particular problem. It turns out that the second derivatives are just too algebraically cumbersome to be of much practical use. Investigating extreme values from first principles is a useful skill worth acquiring. First of all let us solve the equations $f_x = 0, f_y = 0$ to find the stationary values:

$$f_x = ye^{-xy} - xy^2e^{-xy} = (1 - xy)ye^{-xy}$$
$$f_y = (1 - xy)xe^{-xy}$$

The extreme values are thus (0, 0) and all points on the rectangular hyperbola $xy = 1$. First of all, let us investigate the origin using first principles. Close to the origin we can put $x = \varepsilon\cos\theta$ and $y = \varepsilon\sin\theta$, where $0 < \varepsilon << 1$ and $0 \le \theta < 2\pi$. This gives

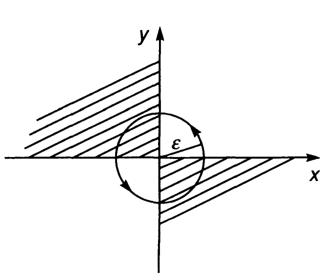


Figure 3.4 The small circle that surrounds the origin.

$$f(x, y) = \varepsilon^2 \cos\theta\sin\theta e^{-\varepsilon^2\cos\theta\sin\theta}$$
$$= \tfrac{1}{2}\varepsilon^2 \sin 2\theta e^{-\varepsilon^2\cos\theta\sin\theta}$$

As θ increases from 0, although the exponential remains positive, $\sin 2\theta = 0$ when $\theta = \frac{\pi}{2}, \pi, \frac{3\pi}{2}, 2\pi$. Thus as the circle of radius ε about the origin is described, $f(x, y)$ is alternately positive and negative giving the classic saddle point pattern as shown in Figure 3.4.

In this figure, the shaded portions denote those parts of the x–y plane where $f(x, y)$ is negative. Now we tackle the more interesting part. All points of the rectangular hyperbola $xy = 1$ are stationary, and an arbitrary point on this hyperbola can be represented by the parametric equations $x = t$, $y = \dfrac{1}{t}$. Thus an arbitrary point on a small circle of radius ε centred at an arbitrary point on this hyperbola has co-ordinates $x = t + \varepsilon \cos\theta$, $y = \dfrac{1}{t} + \varepsilon\sin\theta$. The value of the function $f(x, y)$ at such a point is

$$f = (t + \varepsilon\cos\theta)\left(\frac{1}{t} + \varepsilon\sin\theta\right)e^{-(t+\varepsilon\cos\theta)\left(\frac{1}{t}+\varepsilon\sin\theta\right)}$$

The technique is to expand f in ascending powers of ε utilising the fact that ε is small, retaining only the first non-zero term. Since the point $\left(t, \dfrac{1}{t}\right)$ corresponds to an extreme value of the function f for any t, the $O(\varepsilon)$ term must vanish and we must be prepared to expand to $O(\varepsilon^2)$ at least. To help in the mechanics of this process, computer algebra could (should?) be employed, but for those without access to this, the following gives the details. Note that

$$(t + \varepsilon\cos\theta)\left(\frac{1}{t} + \varepsilon\sin\theta\right) = 1 + \varepsilon\left(t\sin\theta + \frac{1}{t}\cos\theta\right) + \varepsilon^2\cos\theta\sin\theta$$

so that $$e^{-(t+\varepsilon\cos\theta)\left(\frac{1}{t}+\varepsilon\sin\theta\right)} = \frac{1}{e}\,e^{-\varepsilon\left(t\sin\theta+\frac{1}{t}\cos\theta\right)}\,e^{-\varepsilon^2\cos\theta\sin\theta}$$

$$= \frac{1}{e}\left(1 - e\left(t\sin\theta + \frac{1}{t}\cos\theta\right) + \frac{1}{2!}\varepsilon^2\left(t\sin\theta + \frac{1}{t}\cos\theta\right)^2\right)(1 - \varepsilon^2\cos\theta\sin\theta)$$

Hence

$$f = \frac{1}{e}\left(1 + \varepsilon\left(t\sin\theta + \frac{1}{t}\cos\theta\right) + \varepsilon^2\cos\theta\sin\theta\right)$$

$$\times\left(1 - \varepsilon\left(t\sin\theta + \frac{1}{t}\cos\theta\right) + \frac{1}{2}\varepsilon^2\left(t\sin\theta + \frac{1}{t}\cos\theta\right)^2\right)(1 - \varepsilon^2\cos\theta\sin\theta)$$

$$= \frac{1}{e}\left(1 + \varepsilon^2\left(-\frac{1}{2}\left(t\sin\theta + \frac{1}{t}\cos\theta\right)^2\right) + O(\varepsilon^3)\right)$$

(You need your wits about you for this last step!) Now, we have that $f\left(t, \dfrac{1}{t}\right) = \dfrac{1}{e}$ thus $f(x, y) - f\left(t, \dfrac{1}{t}\right) = f - \dfrac{1}{e} < 0$ for all t and for all θ. Thus the hyperbola $xy = 1$ is entirely composed of maxima. An attempt to visualise this rather odd surface has been made in Figure 3.5.

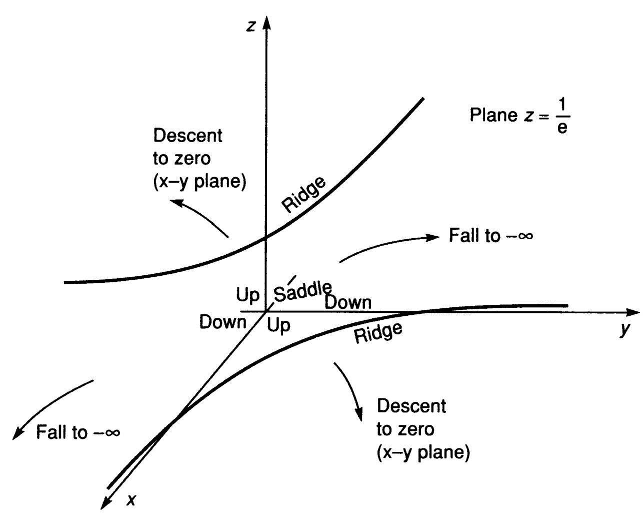


Figure 3.5 The behaviour of the surface $z = xye^{-xy}$ (z is always $\leq \frac{1}{e}$).

Example 3.7 Find an extreme value for the function

$$f(x, y, z) = \frac{e^{x+y+z}}{(e^a + e^x)(e^x + e^y)(e^y + e^z)(e^z + e^b)}$$ and find the value of $f(x, y, z)$ at this point.

Solution For an extreme value, $f_x = f_y = f_z = 0$. To avoid writing too many cumbersome expressions, we note that, owing to the presence of exponentials, f occurs in its own derivative. We can also use symmetry, but with care: Differentiating

$$f_x = f - f\,\frac{e^x}{e^a + e^x} - f\,\frac{e^x}{e^x + e^y} \quad \text{(product rule)}$$

$$f_y = f - f\,\frac{e^y}{e^x + e^y} - f\,\frac{e^y}{e^y + e^z} \quad \text{(product rule)}$$

$$f_z = f - f\,\frac{e^z}{e^y + e^z} - f\,\frac{e^z}{e^z + e^b} \quad \text{(product rule)}$$

Equating each of these to zero and noting that $f \neq 0$ gives

$$\frac{e^x}{e^a + e^x} + \frac{e^x}{e^x + e^y} = 1$$

$$\frac{e^y}{e^x + e^y} + \frac{e^y}{e^y + e^z} = 1$$

$$\frac{e^z}{e^y + e^z} + \frac{e^z}{e^z + e^b} = 1$$

Multiplying each of these three equations by the denominators $(e^a + e^x)(e^x + e^y)$ and rearranging gives

$$\begin{aligned} e^{y+a} &= e^{2x} \\ e^{x+z} &= e^{2y} \\ e^{y+b} &= e^{2z} \end{aligned}$$

from which $y + a = 2x$, $x + z = 2y$ and $y + b = 2z$.

Solving (adding the 1st and 3rd then eliminating $x + z$ using the 2nd is efficient) gives $y = \frac{1}{2}(a+b)$, $x = \frac{3}{4}a + \frac{1}{4}b$, $z = \frac{1}{4}a + \frac{3}{4}b$ as an extreme value. These values are now inserted into the formula for $f(x, y, z)$. We note that, using x_0, y_0, z_0 to denote the extremum:

$$\begin{aligned} e^{x_0+y_0+z_0} &= e^{\frac{3}{2}(a+b)} \\ e^a + e^{x_0} &= e^a\left(1 + e^{\frac{1}{4}(b-a)}\right) \\ e^{x_0} + e^{y_0} &= e^{\frac{3}{4}a+\frac{1}{4}b}\left(1 + e^{\frac{1}{4}(b-a)}\right) \\ e^{y_0} + e^{z_0} &= e^{\frac{1}{2}a+\frac{1}{2}b}\left(1 + e^{\frac{1}{4}(b-a)}\right) \\ e^{z_0} + e^b &= e^{\frac{1}{4}a+\frac{3}{4}b}\left(1 + e^{\frac{1}{4}(b-a)}\right) \end{aligned}$$

The denominator of $f(x_0, y_0, z_0)$ is thus $e^{\frac{5}{2}a+\frac{3}{4}b}\left(1 + e^{\frac{1}{4}(b-a)}\right)^4$, whence

$f(x_0, y_0, z_0) = \dfrac{e^{-a-\frac{1}{2}b}}{\left(1 + e^{\frac{1}{4}(b-a)}\right)^4}$, which can be written $\dfrac{e^{-\frac{1}{2}a-b}}{16\cosh^4 \frac{1}{8}(a-b)}$.

We note that the algebra becomes far too cumbersome to ascertain whether this extremum is a maximum, minimum or whatever by hand. However, computer algebra comes to the rescue and can show that

$$f_{xx} < 0, \quad \begin{vmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{vmatrix} > 0, \quad \text{and} \quad \begin{vmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{yx} & f_{yy} & f_{yz} \\ f_{zx} & f_{zy} & f_{zz} \end{vmatrix} < 0 \text{ at the extremum}$$

$(x, y, z) = (x_0, y_0, z_0)$. Hence $f(x_0, y_0, z_0)$ is a maximum.

Example 3.8 Show that $f(x, y, z)$ has an extremum subject to the condition $g(x, y, z) = 0$ provided $\phi_x = 0$ and $\phi_y = 0$, where $\phi(x, y, z, \lambda) = f(x, y, z) + \lambda g(x, y, z)$.

Solution Since $g(x, y, z) = 0$, we can, in principle, write z as a function of x and y, say $z = h(x, y)$. Substituting this into $f(x, y, z)$ gives $f(x, y, h(x, y))$. Hence for an extremum, using the chain rule:

$$\frac{\mathrm{d}f}{\mathrm{d}x} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial h}\frac{\partial h}{\partial x} = 0 \tag{1}$$

and

$$\frac{\mathrm{d}f}{\mathrm{d}y} = \frac{\partial f}{\partial y} + \frac{\partial f}{\partial h}\frac{\partial h}{\partial y} = 0 \tag{2}$$

However, $g(x, y, z) = 0$, hence we also have the equations

$$\frac{\mathrm{d}g}{\mathrm{d}x} = \frac{\partial g}{\partial x} + \frac{\partial g}{\partial h}\frac{\partial h}{\partial x} = 0 \tag{3}$$

and

$$\frac{\mathrm{d}g}{\mathrm{d}y} = \frac{\partial g}{\partial y} + \frac{\partial g}{\partial h}\frac{\partial h}{\partial y} = 0 \tag{4}$$

Eliminating $\dfrac{\partial h}{\partial x}$ from equations (1) and (3), and $\dfrac{\partial h}{\partial y}$ from equations (2) and (4) leads to $\dfrac{\partial f}{\partial x}\dfrac{\partial g}{\partial h} - \dfrac{\partial g}{\partial x}\dfrac{\partial f}{\partial h} = 0$ and $\dfrac{\partial f}{\partial y}\dfrac{\partial g}{\partial h} - \dfrac{\partial g}{\partial y}\dfrac{\partial f}{\partial h} = 0$. Provided the quantities $\dfrac{\partial f}{\partial h} \neq 0$, $\dfrac{\partial g}{\partial h} \neq 0$, we can divide appropriately and rewrite these two equations as:

$$\frac{\partial f}{\partial x} + \left(-\frac{\partial f}{\partial h}\bigg/\frac{\partial g}{\partial h}\right)\frac{\partial g}{\partial x} = 0$$

and

$$\frac{\partial f}{\partial x} + \left(-\frac{\partial f}{\partial h}\bigg/\frac{\partial g}{\partial h}\right)\frac{\partial g}{\partial y} = 0$$

Writing $\lambda = -\dfrac{\partial f}{\partial h}\bigg/\dfrac{\partial g}{\partial h}$ shows that these two equations are equivalent to $\dfrac{\partial \phi}{\partial x} = 0$ and $\dfrac{\partial \phi}{\partial y} = 0$ where we have written $\phi = f + \lambda g$. This establishes the fundamental result for the technique of undetermined or Lagrange multipliers (λ is the multiplier).

Example 3.9 A rectangular box is just contained within the confines of an ellipsoid which has equation $2x^2 + 3y^2 + z^2 = 18$, with each edge parallel to one of the co-ordinate axes. Find its greatest possible volume.

Solution

This problem is ideally suited to be solved using the Lagrange multiplier technique derived in the last example. Figure 3.6 shows the situation.

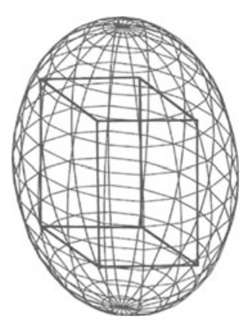


Figure 3.6 An ellipsoid circumscribing a rectangular box.

The sides of the rectangular box (cuboid) are $2x$, $2y$ and $2z$. Hence the volume, which is given by $V = 8xyz$, is the quantity that needs to be maximised, and the constraint is the fact that the corners of the cuboid $\pm x$, $\pm y$, $\pm z$) must lie on the ellipsoid $2x^2 + 3y^2 + z^2 = 18$. We therefore form the function

$$\phi = 8xyz + \lambda(2x^2 + 3y^2 + z^2 - 18)$$

and solve the four equations $\dfrac{\partial \phi}{\partial x} = \dfrac{\partial \phi}{\partial y} = \dfrac{\partial \phi}{\partial z} = \dfrac{\partial \phi}{\partial \lambda} = 0$. These are

$$8yz + 4\lambda x = 0 \qquad (1)$$
$$8xz + 6\lambda x = 0 \qquad (2)$$
$$8xy + 2\lambda z = 0 \qquad (3)$$

and
$$2x^2 + 3y^2 + z^2 - 18 = 0 \qquad (4)$$

The quickest way to solve these is to multiply (1) by x, (2) by y and (3) by z then eliminate $8xyz$ to get

$$\lambda(4x^2 - 6y^2) = 0, \qquad \lambda(6y^2 - 2z^2) = 0$$

Since $\lambda \neq 0$, this implies $4x^2 = 6y^2 = 2z^2$. Substituting this into equation (4) gives

$$z^2 = 6, \quad y^2 = 2, \quad x^2 = 3$$

so that $V = 8\sqrt{3}\sqrt{2}\sqrt{6} = 48$. Second derivatives should now be calculated, but in cases like this where it is reasonably easy to visualise what is happening it is not necessary. $V = 48$ must be a maximum since $V = 0$ if the cuboid is flat (which is certainly a possibility but an obvious minimum), and the saddle point possibility is removed once it is realised that V can never exceed 48.

Example 3.10 Find the shortest distance from the origin to the curve of intersection of the two surfaces $xyz = a$, $y = bx$ (a and b given constants).

Solution This problem is a little more tricky to visualise than the previous example, however it helps to know that the surface $xyz = a$ is asymptotic to all three co-ordinate planes ($x = 0$, $y = 0$, $z = 0$), as if the corner that is the origin has been artistically rounded off. The curve we need to consider is the intersection of this with the plane $y = bx$, which will resemble a hyperbola. Although such a visualisation is useful, the beauty of the Lagrange multiplier technique is that it is not necessary. We simply treat $xyz = a$, $y = bx$ as constraints on the function we wish to minimise which is $r = \sqrt{x^2 + y^2 + z^2}$, the distance of an arbitrary point (x, y, z) from the origin. Note also that the values of x, y and z that minimise r will also minimise r^2 so we can work with the altogether cleaner expression $x^2 + y^2 + z^2$. This is a commonly used and very useful trick that saves considerable algebra. Thus we form

$$\phi(x, y, z, \lambda_1, \lambda_2) = x^2 + y^2 + z^2 + \lambda_1(xyz - a) + \lambda_2(y - bx)$$

where we have used two multipliers λ_1 and λ_2. There is no restriction on the number of multipliers in theory, however in practice too many constraints can lead to the absence of a true extremum. This is the field of constrained optimisation which is largely numerical and outside the scope of this Work Out (but see Chapter 4 for some preliminary ideas). Returning to this example and equating the five first-order partial derivatives of $\phi(x, y, z, \lambda_1, \lambda_2)$ to zero gives the equations

$$2x + \lambda_1 yz - \lambda_2 b = 0 \qquad (1)$$
$$2y + \lambda_1 xz + \lambda_2 = 0 \qquad (2)$$
$$2z + \lambda_1 xy = 0 \qquad (3)$$
$$xyz - a = 0 \qquad (4)$$

and
$$y - bx = 0 \qquad (5)$$

From (3), $\lambda_1 = -\dfrac{2z}{xy}$ whence, using (2), $\lambda_2 = \dfrac{2z^2}{y} - 2y$. Substituting these into (1) yields $2x - \dfrac{2z^2}{x} - \dfrac{2z^2 b}{y} + 2yb = 0$, or $2z^2 = x^2 + y^2$ (call this equation (6)) upon eliminating b via $b = y/x$. The value of $r^2 = x^2 + y^2 + z^2$ is thus $3z^2$ and it remains only to determine z.

Now, from (4) and (5), $z = \dfrac{a}{xy} = \dfrac{a}{bx^2}$ (7)

and, from (6) and (5), $z^2 = \frac{1}{2}(x^2 + y^2) = \frac{1}{2}x^2(1 + b^2)$
Eliminating x^2 between this last equation and (7) gives

$$z = \frac{a(1 + b^2)}{2bz^2}$$

so

$$z^3 = \frac{a(1 + b^2)}{2b}$$

Hence, $r = z\sqrt{3}$ the required minimum distance, is given by $\sqrt{3}\left[\dfrac{a(1 + b^2)}{2b}\right]^{1/3}$. As in the previous example, the species of extremum is self-evident. It must be a minimum as the curve of intersection of the two surfaces goes to infinity, so the distance to any point on this curve from the origin can be as large as we like. For similar reasons the extremum cannot be a saddle point either.

Example 3.11 Find constants a and b such that the definite integral

$$\phi(a, b) = \int_0^{\pi}\left\{\cos x - (ax^2 + b)\right\}^2 dx$$

is a minimum.

Solution Although this problem involves integrals, the method remains unchanged and we equate first derivatives of $\phi(a, b)$ to zero, this time utilising the simplest case of differentiating under the integral sign (constant limits).

$$\int_0^{\pi}\frac{\partial}{\partial a}\left\{\cos x - (ax^2 + b)\right\}^2 dx = 0 \text{ and}$$
$$\int_0^{\pi}\frac{\partial}{\partial b}\left\{\cos x - (ax^2 + b)\right\}^2 dx = 0$$

The equations for a and b are thus

$$a\int_0^{\pi} x^4 dx + b\int_0^{\pi} x^2 dx = \int_0^{\pi} x^2\cos x dx$$
$$a\int_0^{\pi} x^2 dx + b\int_0^{\pi} dx = \int_0^{\pi} \cos x dx$$

where the details of the differentiation have been left to the reader. Integrating (they are all straightforward to do, only $\int_0^{\pi} x^2\cos x dx$ is awkward, requiring integration by parts twice) gives

$$a\frac{\pi^5}{5} + b\frac{\pi^3}{3} = 2\pi$$

$$a\frac{\pi^3}{3} + b\pi = 0$$

and solving these gives $a = -\dfrac{45}{2\pi^4} = -0.231$ and $b = \dfrac{15}{2\pi^2} = 0.760$. What we have done here is to find a specific form of algebraic curve (the parabola $ax^2 + b$) that most closely approximates to $\cos x$. This has been done by minimising the 'square area' between the curves (the integral in the question). It is a form of least squares approximation.

Example 3.12 A rectangular strip of metal of width L is bent up at the sides to form a gutter for rainwater, the cross-section of which is shown in Figure 3.7.

Find the widths of the side pieces (labelled x in Figure 3.7) and the angle through which the sides must be bent (labelled θ in Figure 3.7) in order for the gutter to have a maximum cross-sectional area and hence carry maximum capacity.

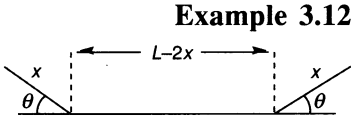


Figure 3.7 The cross-section of the gutter.

Solution This problem is another practical application of finding extrema. The area of the trapezium which forms the cross-sectional area is found using the usual formula

$$\begin{aligned} A &= \tfrac{1}{2}[\text{sum of parallel sides}] \times [\text{distance between them}] \\ &= \tfrac{1}{2}(L - 2x + L - 2x + 2x\cos\theta)(x\sin\theta) \\ &= Lx\sin\theta - 2x^2\sin\theta + \tfrac{1}{2}x^2\sin 2\theta \end{aligned}$$

For an extreme value, $\dfrac{\partial A}{\partial x} = 0$ and $\dfrac{\partial A}{\partial \theta} = 0$ so that we obtain the two equations

$$L\sin\theta - 4x\sin\theta + x\sin 2\theta = 0$$

and

$$Lx\cos\theta - 2x^2\cos\theta + x^2\cos 2\theta = 0$$

These equations look difficult to solve, however, a factor x can be cancelled from the second equation, and writing $\sin 2\theta = 2\sin\theta\cos\theta$ enables a factor of $\sin\theta$ to be cancelled from the first (both $x = 0$ and $\sin\theta = 0$ are obvious minimum values for A). Hence we solve

$$\begin{aligned} &L - 4x + 2x\cos\theta = 0 \\ &L\cos\theta - 2x\cos\theta + x(2\cos^2\theta - 1) = 0 \end{aligned}$$

using $\cos 2\theta = 2\cos^2\theta - 1$. Eliminating $\cos\theta$ from these equations results in the following equation for x:

$$4xL - L^2 - 8x^2 + 2xL + 16x^2 - 8xL + L^2 - 2x^2 = 0$$

which simplifies considerably to $x = \frac{1}{3}L$. Since $\cos\theta = \dfrac{4x - L}{2x}$ this gives $\cos\theta = \frac{1}{2}$ so $\theta = \dfrac{\pi}{3}$.

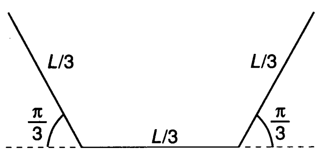


Figure 3.8 The maximum cross-sectional area.

So the drainage channel carries a maximum capacity if the bends are one-third of the total width, and the angle through which the sides are bent is $\frac{\pi}{3}$ (see Figure 3.8).

The value of the cross-sectional area in this maximal case is $\dfrac{L^2\sqrt{3}}{12}$.

Example 3.13 A tea-trolley, shown schematically in Figure 3.9, has a rectangular cross-section of dimensions $p \times q$. It is to be able to round a right-angled bend in a corridor where the width of the corridor is a on one side changing to b on the other. Show that the cross-sectional area is a maximum when $p = \sqrt{a^2 + b^2}$ and $q = \dfrac{ab}{\sqrt{a^2 + b^2}}$.

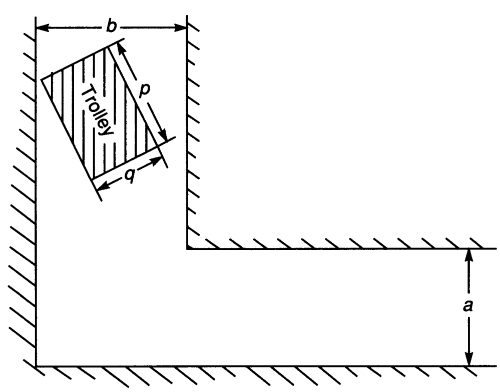


Figure 3.9 The corridor and tea-trolley, a schematic picture.

Solution This is a modelling problem which needs setting up. First of all let us redraw the diagram in critical position as this gives us criteria by which to judge whether the trolley can round the corner without tilting. Figure 3.10 shows the trolley in this position and defines the angle θ as that between the longer side of the trolley and a wall of the corridor.

Using the notation of the figure, we label the longer side of the trolley p and the corner that this

side will just touch (in the limiting case) N. If the longer side is AB then from elementary trigonometry

$$\text{AN} = (b - q\cos\theta)\operatorname{cosec}\theta$$
$$\text{NB} = (a - q\sin\theta)\sec\theta$$

and $$\text{AN} + \text{NB} = \text{AB} = p$$

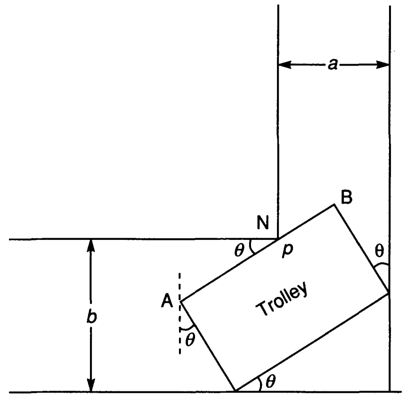


Figure 3.10 The notation used. The trolley is in its limiting position.

Hence, the area S of the trolley $= pq = q\{(b - q\cos\theta)\operatorname{cosec}\theta + (a - q\sin\theta)\sec\theta\}$. In order to achieve a minimum or maximum, we impose the conditions on $S(q,\theta)$:

$\dfrac{\partial S}{\partial q} = 0 \Rightarrow b\operatorname{cosec}\theta - 2q\cos\theta\operatorname{cosec}\theta + a\sec\theta - 2q\sin\theta\sec\theta = 0$, and

$\dfrac{\partial S}{\partial \theta} = 0 \Rightarrow -b\operatorname{cosec}\theta\cot\theta + q\operatorname{cosec}^2\theta + a\sec\theta\tan\theta - q\sec^2\theta = 0$. We solve these two equations for θ and q. Some trigonometric manipulation is required before we arrive at the following simplification:

$$b\cos\theta + a\sin\theta = 2q$$

and $$b\cos^3\theta - a\sin^3\theta = q\cos 2\theta$$

Substituting for q into the expression for $p((b - q\cos\theta)\operatorname{cosec}\theta + (a - q\sin\theta)\sec\theta)$, gives on simplifying

$$p\cos\theta\sin\theta = q$$

and $$q = \tfrac{1}{2}(b\cos\theta + a\sin\theta).$$

At the risk of incurring the wrath of some, more details of the algebra are now given because of the tendency of both human being and computer to go round in circles with all these trigonometric expressions! Using the expression for q just derived, we insert this into $b\cos^3\theta - a\sin^3\theta = q\cos 2\theta$ to obtain

$$\tfrac{1}{2}b\cos\theta(2\cos^2\theta - 1) + \tfrac{1}{2}a\sin\theta(1 - 2\sin^2\theta) = b\cos^3\theta - a\sin^3\theta$$

The cunning step here is the use of different forms of $\cos 2\theta$ in terms of θ which enables the grouping of terms and their subsequent cancellation. The above equation reduces to $-\tfrac{1}{2}b\cos\theta + \tfrac{1}{2}a\sin\theta = 0$

or $$\tan\theta = b/a$$

from which $$\cos\theta = a/\sqrt{a^2 + b^2}$$

It is now straightforward to see that $q = ab/\sqrt{a^2 + b^2}$ and $p = \sqrt{a^2 + b^2}$ by back substitution for θ. Note that the area S derived is the product of the widths of the corridors, and so is obviously a maximum value. To test second derivatives here would be impractical.

3.3 Exercises

3.1. Find the Taylor Series expansions of the functions of two variables given below about the points indicated. In each case, calculate as far as the second-order terms.
(a) e^{xy} about the points (0, 0), (1, 0), (0, 1) and (1, 1),
(b) $\operatorname{cosec}(x + y)$ about the points (1, 0) and (0, 1).
Is there an expansion of $\operatorname{cosec}(x + y)$ about the origin (0, 0)? Give reasons for your answer.

3.2. Determine the first three non-zero terms of the Taylor Series expansion of the function $f(x_1, x_2, x_3, x_4) = \dfrac{1}{2 + x_1x_2x_3x_4}$ about the origin (0, 0, 0, 0).

3.3. Determine the Taylor Series expansion about the general point (u_1, u_2, u_3, u_4) for the function $f(x_1, x_2, x_3, x_4) = \dfrac{1}{x_1} + \dfrac{1}{x_2} + \dfrac{1}{x_3} + \dfrac{1}{x_4}$. (*Hint*: read the remarks in the last half of Example 3.2.) (None of the u's can be zero.)

3.4. Find all the extrema of the functions listed below and classify them:
(a) $f(x, y) = 3x^2 - 2xy + y^2 - 8y$,
(b) $f(x, y) = 4xy - x^4 - y^4$,
(c) $f(x, y) = -xye^{-(x^2+y^2)/2}$,
(d) $f(x, y) = \dfrac{x^2 + 2y^2}{(x + y)^2}$, $\quad x + y \neq 0$,
(e) $f(x, y) = e^x \cos y$.

3.5. Find the extrema of the function $f(x, y) = \sin xy$ in the range $0 \le x, y \le \pi$ and classify them by using the technique introduced in Example 3.6, or otherwise.

3.6. Show that the function $f(x, y) = 4 - \sqrt{x^2 + y^2}$ has a relative maximum at the origin (0, 0), but that none of its partial derivatives exist there. What kind of point is it?

3.7. A function $f(x, y)$ is defined by $f(x, y) = \sin x \sin y$. Investigate all the maxima, minima and saddle points of this function.

3.8. Investigate the extreme values of the function $f(x, y) = x^4 - 2(x - y)^2 + y^4$. In particular, find out in some detail what happens at the origin (0, 0).

3.9. Investigate the function $f(x, y) = \cos\alpha(\cos x + \cos y) - \cos x \cos y$ for possible extrema in the square region $-\dfrac{\pi}{2} \le x, y \le \dfrac{\pi}{2}$, where also $-\dfrac{\pi}{2} \le \alpha \le \dfrac{\pi}{2}$.

3.10. Find the maximum value of the function of three variables $xyz(1 - x - y - z)$.

3.11. Find the extreme value of $f(x, y) = 2x^2 + 4y^2$ subject to the constraint $x^2 + y^2 = 1$.

3.12. The volume of the ellipsoid $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 1$ is $\dfrac{4}{3}\pi abc$.

Show that this is a maximum if $a + b + c = 1$, and that in this case the ellipsoid becomes a sphere.

3.13. A cylindrical tin can, with a top and a bottom, is to be manufactured using 1 m^2 of tin, ignoring waste. What dimensions has the can of maximum volume?

3.14. A rectangular box is to be inscribed in the cone $z = 9 - \sqrt{x^2 + y^2}$, $z \ge 0$. Find the dimensions of the box that maximises its volume.

3.15. Snell's Law for the refraction of light states that $\dfrac{\sin\theta_1}{v_1} = \dfrac{\sin\theta_2}{v_2}$ where θ_1 and θ_2 are the magnitudes of the angles shown in Figure 3.11, and v_1 and v_2 are the velocities of light in the two media.
Use Lagrange multipliers to derive this law using the constraint $x + y = a$ where these are defined in Figure 3.11.

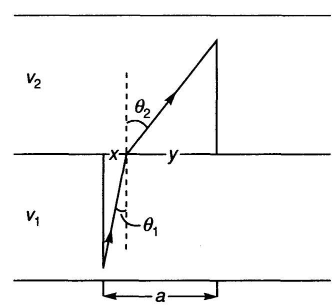


Figure 3.11 Snell's Law of Refraction showing $x + y = a$.

3.16. Reconsider the corridor of Example 3.13. What is the longest ladder that can be taken horizontally around the corner?

Topic Guide

Revision of Matrices
Determinants
Eigenvalues and Eigenvectors
Taylor Polynomials
Newton–Raphson
Steepest Descent
DFP and BFGS Methods

4 Optimisation

4.1 Fact Sheet

An $m \times n$ *matrix* is a rectangular array of numbers with m *rows* and n *columns* and is usually written in the form

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

If $m = n$ the matrix is called square. If $n = 1$ the matrix is called a column vector (or just a vector). The matrices in this chapter (and the whole book) are by and large square. Every square matrix has a number associated with it called its *determinant* which is written

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

A 2×2 determinant is evaluated as follows:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

A 3×3 determinant can be expanded and hence evaluated as follows

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

but this expansion has been made using the top row of the original and, in fact, expansion can be made about *any* row or column of the original 3×3 determinant. The rules are: (a) each element in the chosen row or column is multiplied by the 2×2 determinant of elements *not* in the same row or the same column, and (b), the sign convention shown below is followed in respect of whether there is a + sign or a − sign in front of each term:

$$\begin{vmatrix} + & - & + \\ - & + & - \\ + & - & + \end{vmatrix}$$

In order to evaluate larger-order determinants, such as 4×4 etc., the same rules apply. That is, an $n \times n$ determinant is expanded in terms of $(n - 1) \times (n - 1)$ determinants and so on until the 2×2 stage is reached. The sign convention above is generalised straightforwardly as follows:

$$\begin{vmatrix} + & - & + & \dots & \dots & \dots \\ - & + & - & \dots & & \\ + & - & + & \dots & & \\ \vdots & \vdots & \vdots & & & \\ \vdots & & & & & \end{vmatrix}$$

In practice, we very seldom evaluate any determinant greater than 3×3, the evaluation of larger determinants is particularly well suited to a computational 'DO' loop structure.

Eigenvalues and Eigenvectors

If $\mathbf{A}$ is an $n \times n$ square matrix, and

$$I_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} \text{ is the } n \times n \text{ unit matrix}$$

then the roots of the equation

$$|\mathbf{A} - \lambda I_n| = \begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{vmatrix} = 0$$

defines the *eigenvalues* of the matrix $\mathbf{A}$. If $\lambda^{(r)}$ is one of the eigenvalues of $\mathbf{A}$, then a vector

$$\mathbf{x}^{(r)} = \begin{pmatrix} x_1^{(r)} \\ x_2^{(r)} \\ \vdots \\ x_n^{(r)} \end{pmatrix} \text{ such that } \mathbf{A}\mathbf{x}^{(r)} = \lambda^{(r)}\mathbf{x}^{(r)}$$

is called the *associated eigenvector* $\mathbf{x}^{(r)}$ of the eigenvalue $\lambda^{(r)}$.

Optimisation Techniques

In order to find the maximum or minimum of a complicated function of many variables, numerical techniques have to be applied. Only those that involve the use of partial differentiation and are a direct extension of the techniques introduced in the last two chapters will be examined here. Specialist books on optimisation should be consulted for a complete discussion. Specifically, the techniques covered here are the generalisations of the Newton–Raphson method for finding the roots of an equation, namely the Davidon–Fletcher–Powell method and the less well known Broyden–Fletcher–Goldfarb–Shanno method. Other methods, principally based on discrete as opposed to continuous mathematics, are not found here but form a part of the subject, a very important and extensive subject, called operational research.

The starting point for these methods is the second-order *Taylor polynomial*. This is defined as follows:

$$f(\mathbf{x}) \approx f(\alpha) + \delta\alpha^T\nabla f(\alpha) + \tfrac{1}{2}\delta\alpha^T\mathbf{G}\delta\alpha$$

where the vector $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ represents the n variables in general position in n-dimensional space, whereas the vector $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_n)^T$ represents a specific point and it is implied that $\boldsymbol{\alpha}$ is close to $\mathbf{x}$ in some sense. The superscript 'T' denotes the *transpose* of the matrix. This is defined as follows: A^T, the transpose of A, is the matrix where the rows of A are the columns of A^T and the columns of A are the rows of A^T. Here, the matrix is a vector so the column is conveniently written as the transpose of a row vector. The Taylor polynomial as defined here is the first three terms of the n-variable

Taylor Series. The vector $\nabla f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right)$ is evaluated at $\mathbf{x} = \boldsymbol{\alpha}$, $\mathbf{G}$ is the *Hessian matrix* defined by

$$\mathbf{G} = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$$

again evaluated at $\mathbf{x} = \boldsymbol{\alpha}$ and finally

$$\delta\alpha = (\delta\alpha_1, \delta\alpha_2, \ldots, \delta\alpha_n)^T$$

is $\mathbf{x} - \boldsymbol{\alpha}$ the small n-dimensional vector difference between $\mathbf{x}$ and $\boldsymbol{\alpha}$.

The *Newton–Raphson method* for solving the equation $\nabla f(\mathbf{x}) = \mathbf{0}$ is given by the iteration formula

$$\mathbf{x}^{(r+1)} = \mathbf{x}^{(r)} - [\mathbf{G}(\mathbf{x}^{(r)})]^{-1} \nabla f(\mathbf{x}^{(r)})$$

One version of the method employs line searches where a line in n space is $\mathbf{x} = \mathbf{x}^{(r)} + t\mathbf{s}^{(r)}$ (t is a parameter, see Chapter 5). Defining $\mathbf{H}^{(r)}$ as an approximation to $[\mathbf{G}(\mathbf{x}^{(r)})]^{-1}$ means that $\mathbf{H}^{(r+1)}$ is a better approximation where

$$\mathbf{H}^{(r+1)} = \mathbf{H}^{(r)} + \mathbf{E}^{(r)}$$

and $\mathbf{E}^{(r)}$ is a correction term. If a linear approximation to $\nabla f(\mathbf{x})$ is written

$$\nabla f(\mathbf{x}) = \nabla f(\mathbf{x}^{(r)}) + \mathbf{G}(\mathbf{x}^{(r)})(\mathbf{x} - \mathbf{x}^{(r)})$$

then $\mathbf{E}^{(r)}$, the correction term, takes the form $a\mathbf{u}\mathbf{u}^T$ where a and $\mathbf{u}$ are approximated by terms evaluated at $\mathbf{x}^{(r)}$. This is called a *rank 1* method ($\mathbf{E}^{(r)}$ has one linearly independent row).

If, on the other hand, $\mathbf{E}^{(r)}$ is of the form $a\mathbf{u}\mathbf{u}^T + b\mathbf{v}\mathbf{v}^T$ then it will have *two* linearly independent rows and the method is a *rank 2* method. Two such methods are the *Davidon–Fletcher–Powell* (DFP) method and the *Broyden–Fletcher–Goldfarb–Shanno* (BFGS) method. The DFP method will now be summarised. Written as an algorithm, it looks rather involved. It is advisable to treat what is written below as reference material and to turn straightaway to Examples 4.7 to 4.9 if you want actually to learn how the method works. If one starts with the iterative scheme:

$$\mathbf{x}^{(r+1)} = \mathbf{x}^{(r)} + \mathbf{H}^{(r)}\,\Delta\mathbf{g}^{(r)} + a\mathbf{u}(\mathbf{u}^T\Delta\mathbf{g}^{(r)}) + b\mathbf{v}(\mathbf{v}^T\Delta\mathbf{g}^{(r)})$$

where $\qquad \Delta\mathbf{g}^{(r)} = \nabla f(\mathbf{x}^{(r+1)}) - \nabla f(\mathbf{x}^{(r)})$

then this satisfies the iteration by setting

$$\mathbf{x}^{(r+1)} = \mathbf{x}^{(r)} + a\mathbf{u}(\mathbf{u}^T\Delta\mathbf{g}^{(r)})$$

and $\qquad \mathbf{H}^{(r)}\Delta\mathbf{g}^{(r)} = -b\mathbf{v}(\mathbf{v}^T\Delta\mathbf{g}^{(r)})$

with $\qquad \mathbf{u} = \mathbf{x}^{(r+1)} - \mathbf{x}^{(r)}, \qquad a = \dfrac{1}{\mathbf{u}^T\Delta\mathbf{g}^{(r)}}$

$$\mathbf{v} = \mathbf{H}^{(r)}\Delta\mathbf{g}^{(r)}, \qquad b = -\frac{1}{\mathbf{v}^T\Delta\mathbf{g}^{(r)}}$$

The BFGS method looks even more complicated, but it has the advantage of being more robust (that is, not so sensitive to inexact search). Here it is.

$$\mathbf{H}^{(r+1)} = \mathbf{H}^{(r)} + \left(1 + \frac{(\mathbf{v}^T\Delta\mathbf{g}^{(r)})}{(\mathbf{u}^T\Delta\mathbf{g}^{(r)})}\right)\frac{\mathbf{u}\mathbf{u}^T}{(\mathbf{u}^T\Delta\mathbf{g}^{(r)})} - \frac{\mathbf{u}\mathbf{v}^T + \mathbf{v}\mathbf{u}^T}{(\mathbf{u}^T\Delta\mathbf{g}^{(r)})}$$

where, once again, $\mathbf{u} = \mathbf{x}^{(r+1)} - \mathbf{x}^{(r)}$ and

$$\mathbf{v} = \mathbf{H}^{(r)}\Delta\mathbf{g}^{(r)} = \mathbf{H}^{(r)}(\nabla f(\mathbf{x}^{(r+1)}) - \nabla f(\mathbf{x}^{(r)}))$$

Constrained Optimisation Constrained optimisation remains as in Chapter 3, however it is worth writing it in terms of the notation developed in this chapter. If the function $f(\mathbf{x})$ is to be maximised or minimised subject to a set of constraints

$$c_1(\mathbf{x}) = 0, \quad c_2(\mathbf{x}) = 0, \quad \ldots, c_n(\mathbf{x}) = 0$$

then one forms the function $L(\mathbf{x}, \lambda)$ called the *Lagrangian* where $L(\mathbf{x}, \lambda) = f(\mathbf{x}) + \lambda_1 c_1 + \lambda_2 c_2 + \ldots + \lambda_n c_n$. One can then use an algorithm (such as the DFP or BFGS method) developed for unconstrained problems on this function of $n + r$ variables.

4.2 Worked Examples

Example 4.1 A function of n variables $f(x) = f(x_1, x_2, \ldots, x_n)$ has a minimum value at a certain point. Show that the direction of steepest descent from an arbitrary point $\mathbf{x}_0 = (x_{01}, x_{02}, \ldots, x_{0n})$ to this minimum is along the n vector $\nabla f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right)$.

Solution Consider a small step of fixed length l from $\mathbf{x}_0$. Let this step be in a direction so that an adjacent point is $\mathbf{x}_0 + \mathbf{h}$, where $\mathbf{h} = (h_1, h_2, \ldots, h_n)$. Since the step length is fixed, we have

$$h_1^2 + h_2^2 + \ldots + h_n^2 = l^2$$

that is

$$|\mathbf{h}| = l$$

The problem can thus be formulated in terms of a Lagrange multiplier λ (see Chapter 3) as follows: Since $\mathbf{x} = \mathbf{x}_0 + \mathbf{h}$, the composite function

$$H(\mathbf{x}, \mathbf{h}, \lambda) = f(\mathbf{x}_0 + \mathbf{h}) + \lambda(\mathbf{h}^2 - l^2)$$

and the condition for a minimum is

$$\frac{\partial H}{\partial h_i} = 0, \qquad i = 1, 2, \ldots, n$$

Hence

$$\frac{\partial f}{\partial h_i} + 2\lambda h_i = 0 \qquad i = 1, 2, \ldots, n$$

Thus

$$\mathbf{h} = (h_1, h_2, \ldots, h_n) = -\frac{1}{2\lambda}\left(\frac{\partial f}{\partial h_1}, \frac{\partial f}{\partial h_2}, \ldots, \frac{\partial f}{\partial h_n}\right)$$

$$= -\frac{1}{2\lambda}\nabla f$$

which shows that the vector $\mathbf{h}$ must be parallel to ∇f at the location $\mathbf{x} = \mathbf{x}_0$. It will be shown (in Chapter 7) that, in three dimensions, the vector ∇f is always perpendicular to the contours of f. That ∇f points to the direction of steepest descent is therefore no surprise.

Example 4.2 Use matrix notation to determine Taylor's Theorem for a function of n variables

$$f(\mathbf{x}) = f(x_1, x_2, \ldots, x_n)$$

Solution In two variables, Taylor's Theorem is

$$\begin{aligned} f(a + h, b + k) &= f(a, b) + \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)f(a, b) + \frac{1}{2!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^2 f(a, b) + \ldots \\ &\quad + \frac{1}{n!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^n f(a,b) + R_n \\ &= f(a, b) + (hf_x + kf_y)\,|_{(x,y)=(a,b)} + \tfrac{1}{2}(h^2 f_{xx} + 2hkf_{xy} + k^2 f_{yy})\,|_{(x,y)=(a,b)} + \ldots \end{aligned}$$

In matrix notation this can be written

$$f(a + h, b + k) = f(a, b) + (h\ k)\begin{pmatrix} f_x \\ f_y \end{pmatrix} + \tfrac{1}{2}(h\ k)\begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix}\begin{pmatrix} h \\ k \end{pmatrix} + \ldots$$

It is now reasonably straightforward to extend this to n variables as follows:

$$\begin{aligned} f(a_1 + h_1, a_2 + h_2, \ldots, a_n + h_n) &= f + (h_1, h_2, \ldots, h_n) \cdot \nabla f + \\ &\quad + \tfrac{1}{2}(h_1\ h_2 \ldots h_n)\mathbf{G}\begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{pmatrix} + \ldots \end{aligned}$$

or, more succinctly in vector notation

$$f(\mathbf{a} + \mathbf{h}) = f(\mathbf{a}) + \mathbf{h}^T \cdot \nabla f\,|_{\mathbf{x}=\mathbf{a}} + \tfrac{1}{2}\mathbf{h}^T \cdot \mathbf{G} \cdot \mathbf{h}\,|_{\mathbf{x}=\mathbf{a}} + \ldots$$

where $\mathbf{G}$ is given by

$$\mathbf{G} = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n} \\ \dfrac{\partial^2 f}{\partial x_2 \partial x_1} & \dfrac{\partial^2 f}{\partial x_2^2} & \cdots & \dfrac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1} & \dfrac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$$

and is called the Hessian matrix of $f(\mathbf{x})$.

Example 4.3 Use a method of steepest descent to determine an approximation to the minimum of the function $f(x_1, x_2, x_3) = 2(x_1 - 0.2)^4 + (x_2 - 0.1)^4 + 0.5(x_3 - 0.3)^4$.

Solution Now the minimum value of $f(x_1, x_2, x_3)$ is quite obviously at the point where $x_1 = 0.2$, $x_2 = 0.1$, $x_3 = 0.3$ since f is zero here and positive everywhere else. However, it serves as a good illustration for a search method that incorporates the idea that the direction towards a minimum from anywhere must be in the general direction of the line along which the gradient is largest. (The quickest way to the valley is down the line of steepest slope, as any hiker knows.) However, in general, this line is not straight, so we have to stop and check the direction at each stage. For the algorithm we use here, we stop when the direction of original steepest descent is at right angles to the new (correct) direction of steepest descent. We then head off in this new direction, and so on until the minimum is reached.

Using $f(x_1, x_2, x_3) = 2(x_1 - 0.2)^4 + (x_2 - 0.1)^4 + 0.5(x_3 - 0.3)^4$, we have $\nabla f = [8(x_1 - 0.2)^3, 4(x_2 - 0.1)^3, 2(x_3 - 0.3)^3]$. Let us arbitrarily (and in fact foolishly since we know the answer!) choose (0,0,0) as the starting point. At (0,0,0), the slope ∇f has the value $(-0.064, -0.004, -0.054)$. The direction of greatest slope from the origin is thus the line

$$\begin{aligned}\mathbf{x}^{(1)} &= (0, 0, 0) - \lambda(-0.064, -0.004, -0.054)\\ &= \lambda(0.064, 0.004, 0.054).\end{aligned}$$

See Chapter 6 for more about the vector equation of lines; it is enough to see that on this line, $x_1^{(1)} = 0.064\lambda$, $x_2^{(2)} = 0.004\lambda$, $x_3^{(1)} = 0.054\lambda$ where λ is thought of as a parameter that describes the line, that is, give λ any value between $-\infty$ and $+\infty$ and this gives the co-ordinates of a particular point on the line. As we travel along this straight line from (0,0,0), the slope of the surface will change, so let us determine the value of λ for which the line becomes perpendicular to ∇f (the greatest slope). On the line $\mathbf{x}^{(1)}$, ∇f takes the value

$$\nabla f = [8(0.064\lambda - 0.2)^3, 4(0.004\lambda - 0.1)^3, 2(0.054\lambda - 0.3)^3]$$

Hence $\mathbf{x}^{(1)}. \nabla f = 0$, the condition that $\mathbf{x}^{(1)}$ is at right angles to ∇f occurs at the value of λ for which

$$0.512(0.064\lambda - 0.2)^3 + 0.016(0.004\lambda - 0.1)^3 + 0.108(0.054\lambda - 0.3)^3 = 0$$

Obviously, it is very difficult to solve this cubic equation analytically. A method such as Newton–Raphson (see Chapter 2), a good calculator or computer algebra reveals that $\lambda \approx 4.3$. This value of λ corresponds to the point

$$\mathbf{x}^{(1)}(4.3) = 4.3(0.064, 0.004, 0.054) = (0.27, 0.017, 0.23)$$

We now turn towards the new ∇f direction and work out the new direction of steepest descent. The new line is

$$\begin{aligned}\mathbf{x}^{(2)} &= (0.27, 0.017, 0.23) - \lambda \nabla f \big|_{\mathbf{x}=\mathbf{x}^{(1)}(4.3)}\\ \text{so } \mathbf{x}^{(2)} &= (0.27, 0.017, 0.23) - \lambda[0.0027, -0.00229, -0.000686]\\ &= (0.27 - 0.0027\lambda, 0.017 + 0.0029\lambda, 0.23 + 0.000686\lambda)\end{aligned}$$

Again, λ is determined from finding the value for which this direction becomes perpendicular to the line of greatest slope, that is $\mathbf{x}^{(2)} . \nabla f = 0$. Because of the way the path zigzags, this method is called the *alternating direction* method. Note however that this again involves solving cubic equations in this example. In other examples, the equation for λ can be even more difficult and perhaps impossible to solve. To avoid the tedious exercise of finding λ exactly, note that a respectable guess is every bit as good since we are using an inexact iterative procedure here. For the present problem, if we 'guess' that $\lambda = 20$, we travel along $\mathbf{x}^{(2)}$ to the point (0.216, 0.0628, 0.244) which is respectably near the true minimum (0.2, 0.1, 0.3).

The message here is that it is all very well slavishly to follow the steepest descent algorithm and get to $\mathbf{x}^{(3)}$ where $\mathbf{x}^{(3)} = \mathbf{x}^{(2)} - \lambda\nabla f\big|_{\mathbf{x}=\mathbf{x}^{(2)}(20)}$, but it is much easier to calculate the new value of $\mathbf{x}^{(n)}$ than it is to find λ. This principle is utilised in the two most successful search methods, the Davidon–Fletcher–Powell method and the Broyden–Fletcher–Goldfarb–Shanno method. (See Examples 4.7 and 4.8.)

Exercise 4.4 Show that the general steepest descent/ascent method for finding the maximum or minimum of a function of n variables $f(x_1, x_2, \ldots, x_n)$ leads to a path identical to the alternating direction method by using the three-variable function $f(x_1, x_2, x_3) = 2(x_1 - 0.2)^4 + (x_2 - 0.1)^4 + 0.5(x_3 - 0.3)^4$ of the last example.

Solution In general, a function of n variables $f(x_1, x_2, \ldots, x_n)$ will have many maxima and minima. Suppose we start at an arbitrarily selected point $\mathbf{a} = (a_1, a_2, \ldots, a_n)$. The idea is to move in a direction that, initially at least, is directly towards the nearest maximum or minimum. By a property of the gradient, this direction is ∇f, that is

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n} \right)$$

Hence, from the point **a** we proceed in the direction dictated by the vector ∇f evaluated at $\mathbf{x} = \mathbf{a}$. Suppose we write $\mathbf{h} = \mathbf{x} - \mathbf{a}$, then we ask the question: how do we determine $\mathbf{h}$? The steepest descent/ascent method suggests that the *step length* $|\mathbf{h}|$ where $|\mathbf{h}|^2 = h_1^2 + h_2^2 + \ldots + h_n^2$ should be a constant. It is then a question of maximising or minimising the difference

$$\Delta f = f(a_1 + h_1, a_2 + h_2, \ldots, a_n + h_n) - f(a_1, a_2, \ldots, a_n)$$

subject to the constraint $h_1^2 + h_2^2 + \ldots + h_n^2 =$ constant. This can be formulated in terms of Lagrange multipliers by defining

$$H(h_1, h_2, \ldots, h_n, \lambda) = \Delta f + \lambda(h_1^2 + h_2^2 + \ldots + h_n^2 - \text{const})$$

and then solving the equations $\dfrac{\partial H}{\partial h_i} = 0$ where $i = 1, 2, \ldots, n$ and also $\dfrac{\partial H}{\partial \lambda} = 0$. This is equivalent to solving $\dfrac{\partial f}{\partial h_i} = -2\lambda h_i$ $i = 1, 2, \ldots, n$. What these equations are demanding is that ∇f is parallel to the direction **h**. This is precisely the direction of steepest descent/ascent (as expected). On this line of steepest descent/ascent we can write $x_i = a_i + \lambda \left.\dfrac{\partial f}{\partial x_i}\right|_{x_i = a_i}$ for $i = 1, 2, \ldots, n$ and so $f(x_1, x_2, \ldots, x_n)$ becomes a function of a single variable λ, say $F(\lambda)$. Setting $\dfrac{\mathrm{d}F}{\mathrm{d}\lambda} = 0$ will find the maximum or minimum values of λ, whence **h** is completely determined. This is, it turns out, the same as the alternating direction method. The proof of this is not given here, but as indicated in the question we can demonstrate the truth of this by running through the calculations with the previous three variable $f(x_1, x_2, \ldots, x_n)$. With $f(x_1, x_2, x_3) = 2(x_1 - 0.2)^4 + (x_2 - 0.1)^4 + 0.5(x_3 - 0.3)^4$, we have already calculated that $\nabla f = [8(x_1 - 0.2)^3, 4(x_2 - 0.1)^3, 2(x_3 - 0.3)^3]$. At the origin, our choice of **a**, we have $f = 0.00735$ and $\nabla f = (0.064, 0.004, 0.054)$. The search direction is thus $x_1 = 0.064\lambda$, $x_2 = 0.004\lambda$, $x_3 = 0.054\lambda$; this is all precisely as in Example 4.3. We now form the function $F(\lambda)$ which is

$$F(\lambda) = 2(0.064\lambda - 0.2)^4 + (0.004\lambda - 0.1)^4 + 0.5(0.054\lambda - 0.3)^4$$

$\dfrac{\mathrm{d}F}{\mathrm{d}\lambda} = 0$ demands that we solve the equation:

$$0.512(0.064\lambda - 0.2)^3 + 0.016(0.004\lambda - 0.1)^3 + 0.108(0.054\lambda - 0.3)^3 = 0$$

which is precisely the same equation as that in Example 4.3. It is worth making a comment here. The steepest descent/ascent method is not absolutely identical to the alternating direction method, but in practice the two become identical because, as an extremum is approached, the steps get smaller and smaller. For very small steps the search directions become parallel to the axes and the perpendicularity of the search direction to the direction of steepest descent/ascent follows from the gradient vector being perpendicular to the contours (see Chapter 7). A different approach to search techniques is given in the next two examples.

Example 4.5 Derive a generalisation of the Newton–Raphson method (see Example 1.11) to solve the set of n equations $f_i(\mathbf{x}) = 0$ $\quad i = 1, 2, \ldots, n$; $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$.

Solution Let us start with the statement of the one-variable Newton–Raphson algorithm of Chapter 1, Example 1.11. The sequence $(y_1, y_2, \ldots, y_r, \ldots)$ where

$$y_{k+1} = y_k - \frac{g(y_k)}{g'(y_k)}$$

defines a closer and closer approximation to the zeros of the function $g(y)$. Here y has been used for the variable name to avoid confusion with the notation used for n-dimensional co-ordinates $(x_1, x_2, \ldots, x_n)$. The Newton–Raphson formula is derived from the first two terms of the Taylor Series, that is

$$g(y + \delta y) = g(y) + \delta y g'(y), \text{ where } \delta y = y_{k+1} - y_k, \ y = y_k.$$

In n dimensions, we follow this and try to deduce a formula using the first two terms of the n-dimensional Taylor Series.

Writing the n-variable Taylor Series as follows:

$$f_1(x_1 + \delta x_1, x_2 + \delta x_2, \ldots, x_n + \delta x_n) = f_1(x_1, x_2, \ldots, x_n) + \delta x_1 \frac{\partial f_1}{\partial x_1} + \delta x_2 \frac{\partial f_1}{\partial x_2} + \ldots + \delta x_n \frac{\partial f_1}{\partial x_n}$$

$$f_2(x_1 + \delta x_1, x_2 + \delta x_2, \ldots, x_n + \delta x_n) = f_2(x_1, x_2, \ldots, x_n) + \delta x_1 \frac{\partial f_2}{\partial x_1} + \delta x_2 \frac{\partial f_2}{\partial x_2} + \ldots + \delta x_n \frac{\partial f_2}{\partial x_n}$$

$$\vdots$$

$$\vdots$$

$$\vdots$$

$$f_n(x_1 + \delta x_1, x_2 + \delta x_2, \ldots, x_n + \delta x_n) = f_n(x_1, x_2, \ldots, x_n) + \delta x_1 \frac{\partial f_n}{\partial x_1} + \delta x_2 \frac{\partial f_n}{\partial x_2} + \ldots + \delta x_n \frac{\partial f_n}{\partial x_n}$$

Using matrix notation, we obtain

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix} \begin{pmatrix} \delta x_1 \\ \delta x_2 \\ \vdots \\ \delta x_n \end{pmatrix} = \begin{pmatrix} f_1(\mathbf{x} + \delta\mathbf{x}) - f_1(\mathbf{x}) \\ f_2(\mathbf{x} + \delta\mathbf{x}) - f_2(\mathbf{x}) \\ \vdots \\ f_n(\mathbf{x} + \delta\mathbf{x}) - f_n(\mathbf{x}) \end{pmatrix}$$

If we assume that $f_i(\mathbf{x} + \delta\mathbf{x})$ can be ignored in comparison with $f_i(\mathbf{x})$ on the grounds that $\mathbf{x} + \delta\mathbf{x}$ is closer to a zero of the general equation $f_i(\mathbf{x}) = 0$ than is $\mathbf{x}$, the right-hand vector is approximately $-\mathbf{f}(\mathbf{x})$. Hence the above matrix equation is written succinctly as

$$\mathbf{F}\delta\mathbf{x} = -\mathbf{f}(\mathbf{x})$$

or, inverting $\mathbf{F}$

$$\delta\mathbf{x} = -\mathbf{F}^{-1}\mathbf{f}(\mathbf{x})$$

which implies

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{F}^{-1}\mathbf{f}(\mathbf{x}_k)$$

where we have labelled $\mathbf{x}$ as $\mathbf{x}_k$ to denote that this is the kth step of an iterative process. (To carry this notation through the entire problem would have meant a double suffix notation which, we believe, would have been very confusing.) This then is the direct generalisation of the Newton–Raphson method.

For extremum problems, this can be utilised directly to give an alternative approach to the steepest descent/ascent methods. At an extrema of the function $f(x_1, x_2, \ldots, x_n)$, we have that $\frac{\partial f}{\partial x_1} = 0$, $\frac{\partial f}{\partial x_2} = 0$, $\frac{\partial f}{\partial x_3} = 0, \ldots, \frac{\partial f}{\partial x_n} = 0$. We thus set $f_i = \frac{\partial f}{\partial x_i}$, $i = 1, 2, \ldots, n$ so that the extremum corresponds to $f_i = 0$ for all i. The matrix $\mathbf{F}$ thus equals the *Hessian* $\mathbf{G}$ of the function f, and the Newton–Raphson scheme

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{G}^{-1}\mathbf{f}(\mathbf{x}_k)$$

thus represents a numerical step-by-step method to find extrema. The limitations of this generalised Newton–Raphson method are of similar character to those of the one-variable method, but such questions of accuracy lie outside the scope of this Work Out.

Example 4.6 Use the two-variable Newton–Raphson technique, alongside steepest descent, to find the minimum of the function $f(\mathbf{x}) = (3x_1 - 2)^2 + (x_2 - 5)^2$ using as start vertex $\mathbf{x}_1 = (1, 1)^T$.

Solution Two points need to be made before we start this problem. First of all, to use the Newton–Raphson method with a function of more than two variables is virtually impractical by hand because of the need to invert matrices at each step. This shows once again that we are brushing against the topic of Numerical Methods. Also, this particular example is for illustrative purposes only, since the values $x_1 = \frac{2}{3}$, $x_2 = 5$ are an obvious solution. The method computes as follows:

$$\mathbf{f}(\mathbf{x}) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}\right)^T = (18x_1 - 12, 2x_2 - 10)^T$$

Also
$$\mathbf{F} = \mathbf{G} = \begin{pmatrix} 18 & 0 \\ 0 & 2 \end{pmatrix}$$

so
$$\mathbf{G}^{-1} = \begin{pmatrix} \frac{1}{18} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}, \text{ and thus}$$

$$\delta\mathbf{x} = \mathbf{G}^{-1}\mathbf{f}(\mathbf{x}_1) = \begin{pmatrix} \frac{1}{18} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}\begin{pmatrix} 6 \\ -8 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} \\ -4 \end{pmatrix}$$

This gives the direction of the search towards the minimum as $\lambda\begin{pmatrix} \frac{1}{3} \\ -4 \end{pmatrix}$. Hence the minimum of the function is approached by minimising $f(\mathbf{x}_1 + \lambda\delta\mathbf{x})$. Now

$$\mathbf{x}_1 + \lambda\delta\mathbf{x} = (1, 1)^T + \lambda(\tfrac{1}{3}, -4)^T, \text{ so } x_1 = 1 + \tfrac{1}{3}\lambda \text{ and } x_2 = 1 - 4\lambda \text{ and therefore}$$

$$\begin{aligned} f(\mathbf{x}) &= (3(1 + \tfrac{1}{3}\lambda)-2)^2 + (1 - 4\lambda - 5)^2 \\ &= (1 + \lambda)^2 + (4 + 4\lambda)^2 \end{aligned}$$

Letting $\dfrac{\mathrm{d}f}{\mathrm{d}\lambda} = 0$ gives the equation $2(1 + \lambda) + 8(4 + 4\lambda) = 0$ or $\lambda = -1$; whence $x_2 = (1 - \frac{1}{3}, 1 + 4)^T = (\frac{2}{3}, 5)^T$. This is of course the exact minimum, and it is generally true that the Newton–Raphson technique is exact for quadratic functions. (This is because the derivative of a quadratic is linear, therefore the direction derived must point right at the true minimum of f.)

Example 4.7 Use the function $f(x_1, x_2) = x_1^2 - x_1x_2 + 3x_2^2$ to demonstrate the Davidon–Fletcher–Powell (DFP) method for finding the minimum of a function of many variables.

Solution The DFP method is a practical method for locating a maximum or minimum of a given function of many variables. It is based on the Newton–Raphson method but avoids the necessity of finding the inverse of the Hessian matrix $\mathbf{G}$ at each step. The DFP method is written schematically as $f(\mathbf{x}_{i+1}) = f(\mathbf{x}_i) - k\mathbf{H}_i\mathbf{g}_i$ where $\mathbf{g}_i = \nabla\mathbf{f}_i$, $\mathbf{H}_i$ is a matrix (in some sense close to $\mathbf{G}^{-1}$), $\mathbf{x}_i$ is the approximation to the extremum of $f(\mathbf{x})$ and k is a constant. The DFP method is one of a class called *quasi Newton methods*. In fact, the previous two schemes can be included as special cases of the scheme $f(\mathbf{x}_{i+1}) = f(\mathbf{x}_i) - k\mathbf{H}_i\mathbf{g}_i$ as follows: If $k = \lambda_{\min}$, and $\mathbf{H}_i = \mathbf{I}$, the unit matrix, the scheme is the method of steepest descent. If, on the other hand, $k = 1$ and $\mathbf{H}_i = \mathbf{G}_i^{-1}$ we regain the Newton–Raphson method. In the first case, the method can often be inefficient, and the second method is often impractical. The DFP method gets around both of these difficulties as follows: first set $\mathbf{H}_0 = \mathbf{I}$ the unit matrix, then $\mathbf{H}$ is updated via the formula

$$\mathbf{H}_{i+1} = \mathbf{H}_i - \frac{\mathbf{H}_i\mathbf{y}_i\mathbf{y}_i^T\mathbf{H}_i}{\mathbf{y}_i^T\mathbf{H}_i\mathbf{y}_i} + \frac{\mathbf{h}_i\mathbf{h}_i^T}{\mathbf{h}_i^T\mathbf{y}_i}$$

where $\mathbf{h}_i = \mathbf{x}_{i+1} - \mathbf{x}_i$ is the step length and $\mathbf{y}_i = \nabla f_{i+1} - \nabla f_i$. This procedure is carried out numerically, and for the quadratic function given, $f(x_1, x_2) = x_1^2 - x_1x_2 + 3x_2^2$ terminates in only two iterations.

The start vertex is (1,1), that is $x_1 = 1$, $x_2 = 1$, so $f = f(1,1) = 3$ and $\nabla f = [2x_1 - x_2, -x_1 + 6x_2]^T$ so at the point (1,1), $\nabla f = [1,5]^T$. The first step is thus

$$\mathbf{x}_1 = \begin{pmatrix}1\\1\end{pmatrix} + \lambda\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}\begin{pmatrix}1\\5\end{pmatrix} \text{ so}$$

$$x_{11} = 1 + \lambda$$

$$x_{12} = 1 + 5\lambda$$

The parameter λ is determined by minimising $f(1 + \lambda, 1 + 5\lambda)$ that is

$$(1 + \lambda)^2 - (1 + \lambda)(1 + 5\lambda) + 3(1 + 5\lambda)^2$$

Approximation methods to determine minimum values of functions are the province of numerical methods, see for example the *Numerical Analysis Work Out* by Peter Turner, but any method can be used to determine the minimum value of this quadratic: it is $\lambda = -0.167$. This gives the modified values

$$\mathbf{x}_1 = \begin{pmatrix}0.833\\-4.843\end{pmatrix} \quad \text{and} \quad \nabla f_1 = \begin{pmatrix}1.501\\0.157\end{pmatrix}$$

To find $\mathbf{H}_1$ we utilise the formula for $\mathbf{H}_{i+1}$ where $\mathbf{h}_0 = \mathbf{x}_1 - \mathbf{x}_0 = \begin{pmatrix}-0.167\\-0.835\end{pmatrix}$ and

$$\mathbf{y}_0 = \nabla f_1 - \nabla f_0 = \begin{pmatrix}0.501\\-4.843\end{pmatrix} \text{ and so}$$

$$\mathbf{H}_1 = \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix} - \frac{1}{23.706}\begin{pmatrix}0.251 & -2.426\\-2.426 & 23.455\end{pmatrix} + \frac{1}{3.96}\begin{pmatrix}0.0279 & 0.139\\0.139 & 0.697\end{pmatrix}$$

$$= \begin{pmatrix}0.996 & 0.137\\0.137 & 0.187\end{pmatrix}$$

$$\mathbf{H}_1\,(\nabla f)_1 = \begin{pmatrix}0.996 & 0.137\\0.137 & 0.996\end{pmatrix}\begin{pmatrix}1.501\\0.157\end{pmatrix} = \begin{pmatrix}1.517\\0.235\end{pmatrix}$$

So next we consider $f(0.833 - \lambda 1.517, 0.165 - \lambda 0.235)$ in order to find the value of λ for which f is a minimum. This value is $\lambda = 0.548$. To calculate the next value of $\mathbf{x}$, $\mathbf{x}_2$ we use the formula $\mathbf{x}_2 = \mathbf{x}_1 - 0.548\begin{pmatrix}1.517\\0.235\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}$ within rounding error.

The DFP method has thus found the minimum $\begin{pmatrix}0\\0\end{pmatrix}$ in two steps. In fact it can be proved that it is indeed always the case that this method will find the minimum (or maximum) of a quadratic function in just two steps for the same reason as mentioned in Example 4.6. In more complicated problems, numerical procedures are required to find the values of λ at each step (the so-called cubic algorithm is a popular choice, see the *Numerical Analysis Work Out* by Peter Turner). The important point however is that only the *first* derivatives of $f(x_1, x_2, \ldots, x_n)$ are required to carry out the method.

Example 4.8 Rework the previous example using the Broyden–Fletcher–Goldfarb–Shanno (BFGS) method.

Solution Many of the details will be omitted since the only distinction between the DFP and BFSG methods is the choice of $\mathbf{H}$. The BFSG and DFP methods both select $\mathbf{H}_0 = \mathbf{I}$, but $\mathbf{H}_1$ is given by the more complex formula $\mathbf{H}_1 = \mathbf{H}_0 + \left(1 + \frac{\mathbf{y}_0^T\mathbf{y}_0}{\mathbf{h}_0^T\mathbf{y}_0}\right)\frac{\mathbf{h}_0\mathbf{h}_0^T}{\mathbf{h}_0^T\mathbf{y}_0} - \frac{\mathbf{h}_0\mathbf{y}_0^T + \mathbf{y}_0\mathbf{h}_0^T}{\mathbf{h}_0^T\mathbf{y}_0}$ where $\mathbf{h}_0$ and $\mathbf{y}_0$ retain their meanings from the previous example. Inserting the values from the previous example (all the calculations that precede this step including the calculation of λ remain the same), that is

$$\mathbf{h}_0 = \begin{pmatrix}-0.167\\-0.835\end{pmatrix} \text{ and } \mathbf{y}_0 = \begin{pmatrix}0.501\\-4.843\end{pmatrix}, \text{ gives}$$

$$\mathbf{H}_1 = \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix} + 1.764\begin{pmatrix}0.0279 & 0.1394\\0.1394 & 0.6972\end{pmatrix} - \frac{1}{3.96}\begin{pmatrix}-0.1673 & 0.3905\\0.3905 & 8.0878\end{pmatrix}$$

$$= \begin{pmatrix} 0.909 & 0.246 \\ 0.246 & 0.187 \end{pmatrix}$$

so

$$\mathbf{H}_1(\nabla f)_1 = \begin{pmatrix} 0.909 & 0.236 \\ 0.246 & 0.187 \end{pmatrix}\begin{pmatrix} 1.501 \\ 0.157 \end{pmatrix} = \begin{pmatrix} 1.403 \\ 0.399 \end{pmatrix}$$

We find, as before, the value of λ that minimises $f(0.833 - \lambda 1.403, 0.165 - \lambda 0.399)$ and this is $\lambda = 0.575$, whence the next approximation to our minimum is

$$\mathbf{x}_2 = \mathbf{x}_1 - \lambda\begin{pmatrix} 1.403 \\ 0.399 \end{pmatrix} = \begin{pmatrix} 0.833 \\ 0.399 \end{pmatrix} - 0.575\begin{pmatrix} 1.403 \\ 0.399 \end{pmatrix} = \begin{pmatrix} 0.03 \\ 0.06 \end{pmatrix}$$

This is not quite $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$, but still well within rounding error. The BFGS method is marginally less accurate for this problem but in general has the property of being more robust than the DFP method for very large problems.

Example 4.9 Consider the function

$$f(x_1, x_2, x_3) = 3x_1^4 - 2x_1^2x_2^2 + 6x_2^2 + 10x_1x_3 + 12x_1x_2 - 5x_3 + 3x_3^2$$

subject to the two constraints $c_1(x_1, x_2, x_3) = x_1^2 + 3x_2 - 10$ and $c_2(x_1, x_2, x_3) = 3x_1 - 5x_3 - 8$. Use both the Davidon–Fletcher–Powell and Broyden–Fletcher–Goldfarb–Shanno methods to find an estimate of the maximum value of f. (Numerical software is necessary to solve this problem.)

Solution The underlying strategy for solving constrained optimisation problems is similar in single- or many-variable problems. The technique used is to form the function, sometimes called the Lagrangian function:

$$L(x_1, x_2, x_3, \lambda_1, \lambda_2) = 3x_1^4 - 2x_1^2x_2^2 + 6x_2^2 + 10x_1x_3 + 12x_1x_2 - 5x_3 + 3x_3^2 \\ + \lambda_1(x_1^2 + 3x_2 - 10) + \lambda_2(3x_1 - 5x_3 - 8)$$

and use the Lagrange multiplier technique (see Chapter 3). In order to proceed, we now use the DFP and BFGS algorithms on the function L. The detailed algebra and arithmetic are far too complex to be attempted by hand, so we will set up the problem and discuss the results. For example, ∇L is given by the rather horrific looking column vector:

$$\nabla L = \begin{pmatrix} 12x_1^3 - 4x_1x_2^2 + 10x_3 + 12x_2 - 2\lambda_1x_1 - 3\lambda_2 \\ -4x_1^2x_2^2 + 12x_2 + 12x_1 - 3\lambda_1 \\ 10x_1 - 5 + 6x_3 - 5\lambda_2 \\ -x_1^2 - 3x_2 + 10 \\ -3x_1 + 5x_3 + 8 \end{pmatrix}$$

and we proceed as before. There are many software packages geared up to solve such problems; each one will ask for a *tolerance*, that is an error bound on the numerically derived values of the parameters ($\mathbf{x}$, L, ∇L, etc.). The value 10^{-5} has been selected here. For the algorithm chosen, using the start conditions $\lambda_1 = \lambda_2 = 0$, $x_1 = x_2 = x_3 = 1$, the DFP method converged after nine iterations to the result $f(x_1, x_2, x_3) = 66.1416$ with $x_1 = 1.4757$, $x_2 = 2.6074$ and $x_3 = -0.7146$ to four decimal places.

The BFGS method, even more labour intensive (on the computer that is) using the same start conditions and the same tolerance 10^{-5} gives virtually the same results also after nine steps. (In fact $f(x_1, x_2, x_3) = 66.1415$, a difference of 0.0001 but this is not significant.) Any further assay into optimisation problems is not appropriate to a text where the principal interest is in the methods of calculus.

4.3 Exercises

(* Those exercises marked with an asterisk can only be done properly if you have access to computer software or a good calculator (programmable)).

4.1. Apply the method of steepest descent to the function $f(x, y) = 4x^2 - 4xy + 2y^2$ starting at the point (2, 3). Perform *two* iterations and determine the value of f after the second iteration.

4.2. Find the maximum value of the function $f(x, y) = -(x - 1)^4 - (x - y)^2$ using *three* iterations of the method of steepest descent.

4.3. Use the first-order Taylor polynomial estimate of ∇f about $\mathbf{x}^{(r)}$, the current estimate of the minimum value of $\mathbf{x}$, to show that the Newton–Raphson method for finding an improved estimate $\mathbf{x}^{(r+1)}$ is $\mathbf{x}^{(r)} - G^{-1}(\mathbf{x}^{(r)})\nabla f(\mathbf{x}^{(r)})$ where G is the Hessian matrix for f.

4.4. Perform *one* iteration of the Newton–Raphson method on the function $f(x, y) = x^4 - 2xy + (y + 2)^2$ starting at the point $(\frac{1}{2}, 1)$. Indicate briefly why there is little point in progressing with this method.

4.5.* Determine the minimum value of the function in Example 4.4 by using enough iterations of the alternating (or line search) method to be accurate to three decimal places.

4.6. Use the Newton–Raphson method three times to improve the guess $x = 1$, $y = -2$ to the minimum of the quartic $f(x, y) = x^4 + xy + y^2 + 2y$.

4.7. Find the minimum value of the function

$$f(x, y, z) = x^2 + 8y^2 + z^2 + 2xy + 4yz + 8y - 2z$$

using the Newton–Raphson method with start values $x = y = z = 0$. Hence demonstrate that it is exact in one iteration for quadratic functions.

4.8. Use the Davidon–Fletcher–Powell method (use *two* cycles) to find the minimum of the function $f(x, y) = (x - y)^2 + \frac{1}{16}(x + y + 1)^4$. Recompute using two steps of each of the steepest descent and the Newton–Raphson method.

4.9.* In an experiment, the following data points (x, y) were obtained: (0,4), (1,2), (3,1), (4,0.5), (5,0.3). We wish to fit an exponential curve of the general form $y = ae^{bx}$ to this data, and we do this by finding the *residual*

$$\mathbf{r} = \begin{bmatrix} a - 4 \\ ae^{b} - 2 \\ ae^{3b} - 1 \\ ae^{4b} - 0.5 \\ ae^{5b} - 0.3 \end{bmatrix}$$

and minimising

$$\mathbf{r}^T\mathbf{r} = (a - 4)^2 + (ae^{b} - 2)^2 + (ae^{3b} - 1)^2 + (ae^{4b} - 0.5)^2 + (ae^{5b} - 0.3)^2$$

(a least squares method). Use the DFP method to minimise this square residual. Show that there is convergence with six decimal place accuracy after four steps of the method.

4.10. Figure 4.1 shows the position of three drainpipes that have

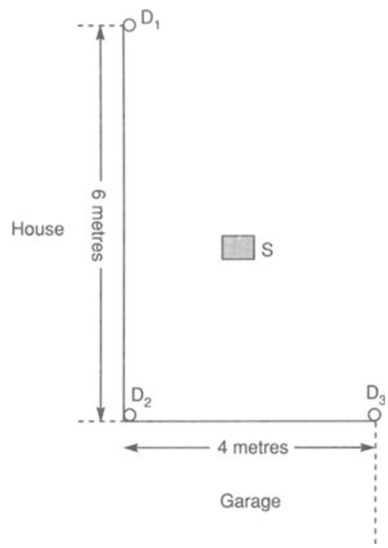


Figure 4.1 A soakaway S, three drains D_1, D_2 and D_3, the house and garage.

to drain the run-off from gutters into a mutual soakaway via three straight drains.

Formulate (but do not solve) as an unconstrained minimisation problem. Outline the procedure for using either the DFP or BFGS algorithm to solve this problem.

4.11.* This question is about Rosenbrock's Function

$$f(x, y) = 100(y - x^2)^2 + (1 - x)^2$$

Starting from the point (−1.2, 1) show that neither the Newton–Raphson nor the steepest descent method works efficiently for this function. Use the DFP method on this function and hence show that convergence takes place to five decimal place accuracy in about 14 iterations to the true minimum (1, 1).

4.12.* Three further 'pathological' functions are:
(a) Powell's Function

$$f(\mathbf{x}) = (x_1 + 10x_2)^2 - 5(x_3 - x_4)^2 + (x_2 - 2x_3)^4 + 10(x_1 - x_4)^4$$ starting with the values $\mathbf{x}^{(0)} = (3, -1, 0, 1)$.

(b) Fletcher and Powell's helix function

$$f(\mathbf{x}) = 100\left((x_3 - 10\theta)^2 + (\sqrt{x_1^2 + x_2^2} - 1)^2\right) + x_3^2$$

$$\text{where } \theta = \begin{cases} \dfrac{1}{2\pi}\tan^{-1}\left(\dfrac{x_2}{x_1}\right) & \text{for } x_1 > 0 \\[2ex] \dfrac{1}{2\pi}\left(\pi + \tan^{-1}\left(\dfrac{x_2}{x_1}\right)\right) & \text{for } x_1 \le 0 \end{cases}$$

starting with the values $\mathbf{x}^{(0)} = (-1, 0, 0)$.
(c) Wood's function

$$f(\mathbf{x}) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2 + 90(x_4 - x_3^2)^2 + (1 - x_3)^2 + 10.1((x_2 - 1)^2 + (x_4 - 1)^2) + 19.8(x_2 - 1)(x_4 - 1)$$

starting with the values $\mathbf{x}^{(0)} = (-3, -1, -3, -1)$.

Try both the DFP and BFGS methods on these functions. If you try either the steepest descent or Newton–Raphson methods, then convergence is either non-existent or painfully slow.

4.13. A function $f(x, y) = x^2 + y^2$ is subject to the constraint $xy = 3$. Form the Lagrangian and find the Hessian matrix for this problem. Hence find the minimum value of $f(x, y)$.

4.14. Show that the point $(\frac{1}{2}, \frac{1}{3}, \frac{7}{4})$ is a stationary point of the function

$$f(x, y, z) = \tfrac{1}{4}x(x + 1) + 6y^2 - 4yz + \tfrac{5}{6}z$$

when subject to the constraints $x^2 + z - 2 = 0$ and $x + 3y + 2z - 5 = 0$. What species of stationary point is it?

4.15.* Use the DFP method to solve the following problem:

Minimise $\quad f(x, y) = x^2 - 2xy + 2y^2 + 6x + 7y$

subject to $\quad x^2 + y^2 \leq 25, \qquad x + 3y \leq 0.$

(*Hint*: turn the inequalities into equalities by adding u and v, then form the Lagrangian. This problem is an example of using the DFP method for inequality constraints and the use of penalty functions u and v.)

4.16.* Use both the DFP and BFGS methods to solve the following problem:

$$f(x, y, z) = x^2 + 2y^2 + 3z^2$$

subject to $\quad x + 2y + 4z = 12$

and $\quad 2x + y + 3z = 10.$

Check your result by using the Newton–Raphson method and state why the method works so well.

Topic Guide

Definition
Addition
Scalar and Vector Products
Triple Products
Vector Equations

5 Vector Analysis

5.1 Fact Sheet

A *vector* is a quantity that has both magnitude and direction. A *scalar*, on the other hand, is a quantity that has magnitude only. A vector is usually denoted by a bold lower case letter, such as **a**, although upper case letters are sometimes also used. A vector so defined is unique, but only as far as magnitude and direction are concerned. It can still be translated parallel to itself and remain the identical vector. Geometrical applications of vectors often employ the notation $\overrightarrow{PQ}$ to denote the vector representation of the line joining the point P to the point Q (in that order). Such a vector is called *line bound.* If a vector **a** is to represent such a vector, then in addition to direction and magnitude, one point on the line of the vector must also be specified.

The vectors $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$ are used to represent the vectors in the direction of the x, y and z axes respectively. The carat above is a symbol used to denote that each is a vector of magnitude one, however this symbol is sometimes omitted in the case of **i, j** and **k.** Vectors of magnitude one are called *unit* vectors, that is their length is unity. This notation is general, that is $\hat{\mathbf{a}}$ denotes a unit vector. Any vector **a** may be written $a_1\hat{\mathbf{i}} + a_2\hat{\mathbf{j}} + a_3\hat{\mathbf{k}}$. The scalar quantities a_1, a_2 and a_3 are called the *components* of the vector **a**. The notation $|\mathbf{a}| = a$ is used to denote the magnitude or length of the vector **a**. In terms of components, $|\mathbf{a}| = \sqrt{a_1^2 + a_2^2 + a_3^2}$. It follows that the unit vector $\hat{\mathbf{a}}$ is given by

$$\hat{\mathbf{a}} = \frac{\mathbf{a}}{|\mathbf{a}|} = \frac{a_1\hat{\mathbf{i}} + a_2\hat{\mathbf{j}} + a_3\hat{\mathbf{k}}}{\sqrt{a_1^2 + a_2^2 + a_3^2}}$$

The quantities $\frac{a_1}{|\mathbf{a}|}, \frac{a_2}{|\mathbf{a}|}$ and $\frac{a_3}{|\mathbf{a}|}$ are called the *direction cosines* of **a**.

The *scalar product*, $\mathbf{a} \cdot \mathbf{b}$, of two vectors **a** and **b** is defined by

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos\theta = ab\cos\theta$$

where θ is the angle between **a** and **b**. $\mathbf{a} \cdot \mathbf{b}$ is a scalar quantity.

The *vector product*, $\mathbf{a} \times \mathbf{b}$, of two vectors **a** and **b** is defined by

$$\mathbf{a} \times \mathbf{b} = |\mathbf{a}||\mathbf{b}| \sin\theta\hat{\mathbf{n}}$$

where θ is still the angle between **a** and **b**, and $\hat{\mathbf{n}}$ is a unit vector in a direction such that **a**, **b** and $\hat{\mathbf{n}}$ are a right-handed system of vectors (see Figure 5.1).

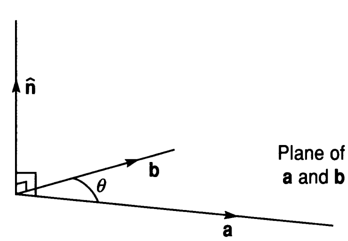


Figure 5.1 The right-handed triad **a**, **b**, $\hat{\mathbf{n}}$.

The *scalar triple product* of three vectors **a**, **b** and **c** is the quantity $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$, and the *vector triple product* of the same three vectors is $\mathbf{a} \times (\mathbf{b} \times \mathbf{c})$. Examples involving these triple products are found from Example 5.18 onwards.

5.2 Worked Examples

Example 5.1 Use a diagram to represent the sum $\mathbf{a} + \mathbf{b}$ where $\mathbf{a} = \hat{\mathbf{i}} + 2\hat{\mathbf{j}}$ and $\mathbf{b} = 2\hat{\mathbf{i}} + \hat{\mathbf{j}}$.

Solution First of all let us draw this diagram, displayed as Figure 5.2.

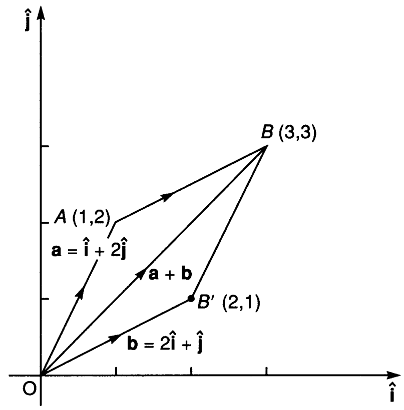


Figure 5.2 This shows **a**, **b** and the sum **a** + **b**.

Algebraically, if $\mathbf{a} = \hat{\mathbf{i}} + 2\hat{\mathbf{j}}$ and $\mathbf{b} = 2\hat{\mathbf{i}} + \hat{\mathbf{j}}$ then $\mathbf{a} + \mathbf{b} = 3\hat{\mathbf{i}} + 3\hat{\mathbf{j}}$. In order to represent this on a graph, first **a** is drawn connecting the origin (0, 0) to the point (1, 2). This line can be thought of as being of length $|\mathbf{a}| = \sqrt{1^2 + 2^2} = \sqrt{5}$ and in the direction of the unit vector $\hat{\mathbf{a}} = \frac{1}{\sqrt{5}}\left(\hat{\mathbf{i}} + 2\hat{\mathbf{j}}\right)$. Of course the vector **a** can represent any line of length $\sqrt{5}$ pointing in this direction, so we are able to move the line OA in Figure 5.2 anywhere parallel to itself. The vector **a** is thus not a *line bound vector*. The vector $\overrightarrow{OA}$, however, is certainly line bound. The second vector **b** may either be drawn as $\overrightarrow{AB}$, on the end of $\overrightarrow{OA}$ as in Figure 5.2, in which case $\mathbf{a} + \mathbf{b}$ is the other side of the triangle $\overrightarrow{OB}$. Or it can be drawn from the origin as $\overrightarrow{OB'}$ in which case $\mathbf{a} + \mathbf{b}$ is the diagonal of the parallelogram which is, of course, once again $\overrightarrow{OB}$. This use of vectors may have been met before as the triangle law of addition (triangle of forces in mechanics) or the parallelogram law of addition (parallelogram of forces in mechanics).

Example 5.2 By examining the vector that joins the origin to the point (5, 7, 9) geometrically, establish that this vector can be represented in component form $5\hat{\mathbf{i}} + 7\hat{\mathbf{j}} + 9\hat{\mathbf{k}}$ where $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$ are unit vectors in the x, y and z directions respectively.

Solution Figure 5.3 shows the vector joining the origin to the point (5,7,9) (point P). This vector, denoted by $\overrightarrow{OP} = \mathbf{r}$ can be derived from the combination $\overrightarrow{OP} = \overrightarrow{OA} + \overrightarrow{AB} + \overrightarrow{BP}$. These are all line bound vectors, and $|\overrightarrow{OA}| = OA = 5$, $|\overrightarrow{AB}| = AB = 7$, $|\overrightarrow{BP}| = BP = 9$. Since $\overrightarrow{OA}$ is in the $\hat{\mathbf{i}}$ direction, $\overrightarrow{AB}$ is in the $\hat{\mathbf{j}}$ direction and $\overrightarrow{BP}$ is in the $\hat{\mathbf{k}}$ direction, and a vector is the product of its magnitude and its unit direction, we have

$$\overrightarrow{OA} = 5\hat{\mathbf{i}},\ \overrightarrow{AB} = 7\hat{\mathbf{j}},\ \overrightarrow{BP} = 9\hat{\mathbf{k}}$$

This means that $\mathbf{r} = \overrightarrow{OP} = 5\hat{\mathbf{i}} + 7\hat{\mathbf{j}} + 9\hat{\mathbf{k}}$ as required. The bold **r** is virtually always the symbol used for the position vector. Another point worth mentioning is that the triple (5, 7, 9) is often used as an alternative to $5\hat{\mathbf{i}} + 7\hat{\mathbf{j}} + 9\hat{\mathbf{k}}$ when it is clear that a vector and not a geometrical point is being represented.

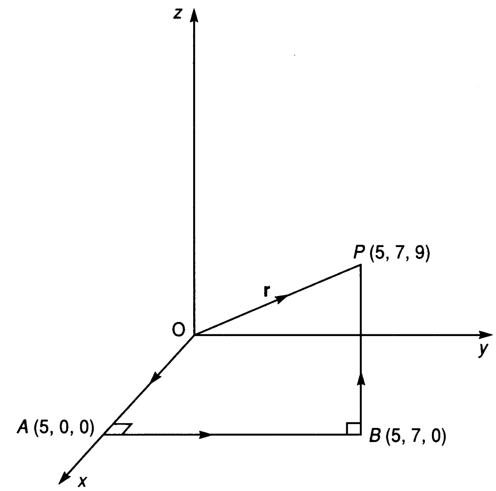


Figure 5.3 The vector $\overrightarrow{OP}$ $= \mathbf{r}$ and its components shown diagrammatically.

Example 5.3 Find the direction cosines of the vector $\mathbf{A} = 3\hat{\mathbf{i}} + 2\hat{\mathbf{j}} + 6\hat{\mathbf{k}}$

Solution The direction cosines of a vector **A** are the cosines of the angles γ_1, γ_2 and γ_3 (shown in Figure 5.4). These are the angles the vector **A** makes with the three co-ordinate axes x, y and z respectively.

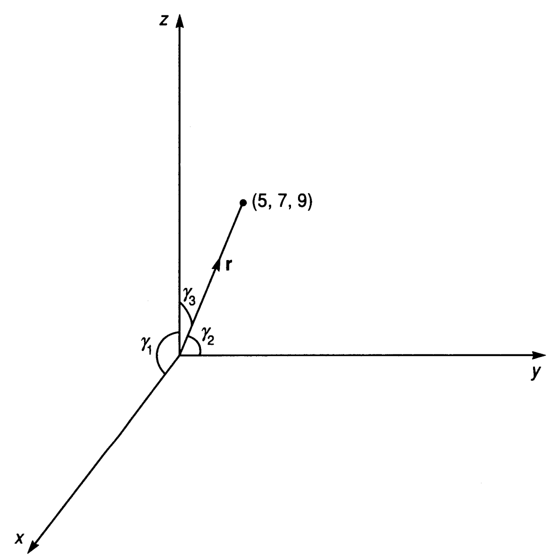


Figure 5.4 The direction cosines of (5, 7, 9) and the cosines of γ_1, γ_2 and γ_3.

A is already given in component form $A_1\hat{\mathbf{i}} + A_2\hat{\mathbf{j}} + A_3\hat{\mathbf{k}}$ whence $A_1 = 3$, $A_2 = 2$ and $A_3 = 6$. These values are the lengths of the projection of **A** on to the co-ordinate axes. They therefore each form one side of the right-angled triangles used to define the direction cosine, the hypotenuse of which is $|\mathbf{A}| = \sqrt{3^2 + 2^2 + 6^2} = \sqrt{49} = 7$. The direction cosines are thus $\frac{3}{7}$, $\frac{2}{7}$ and $\frac{6}{7}$. Of course these could have been found by direct use of the formulae $\frac{A_1}{|\mathbf{a}|}$, $\frac{A_2}{|\mathbf{a}|}$ and $\frac{A_3}{|\mathbf{a}|}$, but the above geometry tells us what they actually are, in particular why they are called cosines.

Example 5.4 Given the three vectors

$$\mathbf{a} = \hat{\mathbf{i}} + \hat{\mathbf{j}} + \hat{\mathbf{k}}$$
$$\mathbf{b} = \hat{\mathbf{i}} - \hat{\mathbf{j}} + \hat{\mathbf{k}}$$
$$\mathbf{c} = \hat{\mathbf{i}} + \hat{\mathbf{j}} - \hat{\mathbf{k}}$$

find (a), the magnitude of $\mathbf{a} + \mathbf{b} + \mathbf{c}$, (b) the unit vector in the direction of $\mathbf{a} - \mathbf{b} + \mathbf{c}$, (c) the direction cosines of $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$, and (d) a unit vector parallel to $\mathbf{a} + \mathbf{c}$.

Solution This problem is purely algebraic and is best treated as an example of manipulating vectors algebraically.

(a) $\mathbf{a} + \mathbf{b} + \mathbf{c} = (\mathbf{i} + \mathbf{j} + \mathbf{k}) + (\mathbf{i} - \mathbf{j} + \mathbf{k}) + (\mathbf{i} + \mathbf{j} - \mathbf{k})$
$= 3\mathbf{i} + \mathbf{j} + \mathbf{k}$ adding components.
Hence the magnitude of $\mathbf{a} + \mathbf{b} + \mathbf{c}$ is $|3\mathbf{i} + \mathbf{j} + \mathbf{k}| = \sqrt{3^2 + 1^2 + 1^2} = \sqrt{11}$.

(b) $\mathbf{a} - \mathbf{b} + \mathbf{c} = (\mathbf{i} + \mathbf{j} + \mathbf{k}) - (\mathbf{i} - \mathbf{j} + \mathbf{k}) + (\mathbf{i} + \mathbf{j} - \mathbf{k})$
$= \mathbf{i} + 3\mathbf{j} - \mathbf{k}$
This is not a line bound vector, hence it represents any vector in the direction $\mathbf{i} + 3\mathbf{j} - \mathbf{k}$. The magnitude of this vector is $\sqrt{1^2 + 3^2 + (-1)^2} = \sqrt{11}$, so a unit vector in this direction is $(\mathbf{i} + 3\mathbf{j} - \mathbf{k})/\sqrt{11}$.

(c) The direction cosines of $(\mathbf{i} + \mathbf{j} + \mathbf{k})$ are $\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}$ and $\frac{1}{\sqrt{3}}$. These are obtained by dividing each of the components of $(\mathbf{i} + \mathbf{j} + \mathbf{k})$ in turn by its magnitude $\sqrt{3}$. Similarly, the direction cosines of $(\mathbf{i} - \mathbf{j} + \mathbf{k})$ are $\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}$ and $\frac{1}{\sqrt{3}}$, and those of $(\mathbf{i} + \mathbf{j} - \mathbf{k})$ are $\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}$ and $-\frac{1}{\sqrt{3}}$.

(d) The vector $\mathbf{a} + \mathbf{c} = 2\mathbf{i} + 2\mathbf{j}$, so a unit vector in this direction is this vector divided by $\sqrt{2^2 + 2^2} = \sqrt{8} = 2\sqrt{2}$, hence a unit vector is $\frac{(\mathbf{i} + \mathbf{j})}{\sqrt{2}}$.

Example 5.5

Using the definition of scalar product, show that, for any vector $\mathbf{a}$

$$\mathbf{a} = (\mathbf{a} \cdot \mathbf{i})\mathbf{i} + (\mathbf{a} \cdot \mathbf{j})\mathbf{j} + (\mathbf{a} \cdot \mathbf{k})\mathbf{k}$$

Solution The scalar product of two vectors is defined as

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos\theta$$

where θ is the angle between the vectors $\mathbf{a}$ and $\mathbf{b}$. Since $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ are unit vectors, and they are mutually perpendicular, we have the following results:

$$\mathbf{i} \cdot \mathbf{i} = 1, \mathbf{j} \cdot \mathbf{j} = 1, \mathbf{k} \cdot \mathbf{k} = 1$$
$$\mathbf{i} \cdot \mathbf{j} = \mathbf{j} \cdot \mathbf{k} = \mathbf{k} \cdot \mathbf{i} = 0$$

Hence for the vectors $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$ and $\mathbf{i}$, we have the scalar product

$$\begin{aligned}\mathbf{a} \cdot \mathbf{i} &= a_1\mathbf{i} \cdot \mathbf{i} + a_2\mathbf{j} \cdot \mathbf{i} + a_3\mathbf{k} \cdot \mathbf{i} \\ &= a_1\end{aligned}$$

Similarly, $\mathbf{a} \cdot \mathbf{j} = a_2$ and $\mathbf{a} \cdot \mathbf{k} = a_3$, whence

$$\begin{aligned}\mathbf{a} &= a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k} \\ &= (\mathbf{a} \cdot \mathbf{i})\mathbf{i} + (\mathbf{a} \cdot \mathbf{j})\mathbf{j} + (\mathbf{a} \cdot \mathbf{k})\mathbf{k}\end{aligned}$$

as required.

Example 5.6

Find a formula for the scalar (or 'dot') product of $\mathbf{a}$ and $\mathbf{b}$ in terms of their components a_1, a_2, a_3, b_1, b_2 and b_3.

Solution Writing $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$ and $\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$ and using the algebra of vectors gives

$$\begin{aligned}\mathbf{a}\cdot\mathbf{b} &= (a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k})\cdot(b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k})\\ &= a_1b_1\mathbf{i}\cdot\mathbf{i} + a_1b_2\mathbf{i}\cdot\mathbf{j} + a_1b_3\mathbf{i}\cdot\mathbf{k}\\ &\quad + a_2b_1\mathbf{j}\cdot\mathbf{i} + a_2b_2\mathbf{j}\cdot\mathbf{j} + a_2b_3\mathbf{j}\cdot\mathbf{k}\\ &\quad + a_3b_1\mathbf{k}\cdot\mathbf{i} + a_3b_2\mathbf{k}\cdot\mathbf{j} + a_3b_3\mathbf{k}\cdot\mathbf{k}\\ &= a_1b_1 + a_2b_2 + a_3b_3\end{aligned}$$

since the unit vectors **i**, **j** and **k** are an orthogonal set (see Example 5.5). The formula

$$\mathbf{a}\cdot\mathbf{b} = a_1b_1 + a_2b_2 + a_3b_3$$

is extremely useful for the computation of the scalar or dot product.

Example 5.7 Find the angle between the following sets of vectors: (a) (1,1,1) and (0,0,1), (b) (1,1,1) and (1,0,1), (c) (1,0,1) and (−1,0,1).

Solution A convenient way of finding the angle between two vectors **a** and **b** is to use the definition of the scalar product

$$\mathbf{a}\cdot\mathbf{b} = |\mathbf{a}||\mathbf{b}|\cos\theta$$

so that

$$\cos\theta = \frac{\mathbf{a}\cdot\mathbf{b}}{|\mathbf{a}||\mathbf{b}|}$$

(a) With $\mathbf{a} = (1,1,1)$ and $\mathbf{b} = (0,0,1)$, we have $\mathbf{a}\cdot\mathbf{b} = 1$

$$|\mathbf{a}| = \sqrt{1^2 + 1^2 + 1^2} = \sqrt{3} \text{ and } |\mathbf{b}| = \sqrt{0^2 + 0^2 + 1^2} = 1$$

Thus $\cos\theta = \dfrac{1}{\sqrt{3}.1} = \dfrac{1}{\sqrt{3}}$ so that $\theta = \cos^{-1}\left(\dfrac{1}{\sqrt{3}}\right) = 0.955$ in radians (or 55 degrees).

(b) With $\mathbf{a} = (1,1,1)$ and $\mathbf{b} = (1,0,1)$, this time $\mathbf{a}\cdot\mathbf{b} = 2$
$|\mathbf{a}| = \sqrt{3}$ as before, but $|\mathbf{b}| = \sqrt{2}$

Thus $\cos\theta = \dfrac{2}{\sqrt{3}\sqrt{2}} = \sqrt{\dfrac{2}{3}}$ so that $\theta = \cos^{-1}\left(\sqrt{\dfrac{2}{3}}\right) = 0.615$ in radians (or 35 degrees).

(c) With $\mathbf{a} = (1,0,1)$ and $\mathbf{b} = (-1,0,1)$, $\mathbf{a}\cdot\mathbf{b} = -1 + 1 = 0$

Therefore, $\cos\theta = 0$ so $\theta = \dfrac{\pi}{2}$ radians or 90 degrees.

If **a** is perpendicular to **b** then $\mathbf{a}\cdot\mathbf{b} = 0$. Likewise, if $\mathbf{a}\cdot\mathbf{b} = 0$ then $\cos\theta = 0$ so $\theta = \dfrac{\pi}{2}$ and **a** must be perpendicular to **b**. We have thus proved that $\mathbf{a}\cdot\mathbf{b} = 0$ if and only if vector **a** is perpendicular to vector **b**. This is another very useful result.

Example 5.8 A tetrahedron is formed from the six vectors **a**, **b**, **c**, **d**, **e** and **f**. If two pairs of opposite edges are perpendicular, show that the third pair is also perpendicular.

Solution The tetrahedron is shown in Figure 5.5. From the triangle law applied to the four faces, we have

$$\mathbf{a} + \mathbf{b} + \mathbf{c} = \mathbf{0}$$
$$\mathbf{a} + \mathbf{f} - \mathbf{e} = \mathbf{0}$$
$$\mathbf{b} - \mathbf{d} + \mathbf{e} = \mathbf{0}$$
$$\mathbf{c} + \mathbf{d} - \mathbf{f} = \mathbf{0}$$

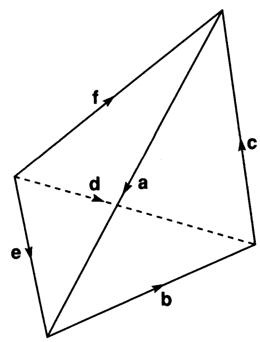


Figure 5.5 The tetrahedron (**d** is behind **a**).

(in fact the fourth equation can be deduced from the first three). Assume that the given pairs of perpendicular edges are **a** and **d**, and **c** and **e**, so $\mathbf{a} \cdot \mathbf{d} = 0$ and $\mathbf{c} \cdot \mathbf{e} = 0$. We thus have to show that $\mathbf{b} \cdot \mathbf{f} = 0$. To do this, we proceed as follows: from the fourth and first of the above equations we have

$$\mathbf{d} = \mathbf{f} - \mathbf{c}$$

$$\mathbf{a} = -\mathbf{b} - \mathbf{c}$$

The third equation is $\mathbf{e} = \mathbf{d} - \mathbf{b}$, and eliminating **d** gives

$$\mathbf{e} = \mathbf{f} - \mathbf{b} - \mathbf{c}$$

We now use the information that $\mathbf{a} \cdot \mathbf{d} = 0$, and $\mathbf{c} \cdot \mathbf{e} = 0$ to give

$$\mathbf{a} \cdot \mathbf{d} = (-\mathbf{b} - \mathbf{c}) \cdot (\mathbf{f} - \mathbf{c}) = -\mathbf{b} \cdot \mathbf{f} - \mathbf{c} \cdot \mathbf{f} + \mathbf{b} \cdot \mathbf{c} + \mathbf{c}^2 = 0$$

and $$\mathbf{c} \cdot \mathbf{e} = \mathbf{c} \cdot (\mathbf{f} - \mathbf{b} - \mathbf{c}) = \mathbf{c} \cdot \mathbf{f} - \mathbf{b} \cdot \mathbf{c} - \mathbf{c}^2 = 0$$

From these last two equations we must have $\mathbf{b} \cdot \mathbf{f} = 0$, the desired result.

Example 5.9 Determine the general equation of a plane and the general equation of a sphere in terms of the position vector $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$.

Solution The general equation of a plane in Cartesian co-ordinates is $a_1x + a_2y + a_3z = c$ where a_1, a_2, a_3 and c are constants. Writing $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$ with **r** as the position vector, this equation can be written as $\mathbf{r} \cdot \mathbf{a} = c$. This may seem only a cosmetic change, but there are two points worth mentioning. First, the equation $\mathbf{r} \cdot \mathbf{a} = c$ represents the locus of points for which $|\mathbf{r}||\mathbf{a}| \cos\theta = c$ (see Figure 5.6)

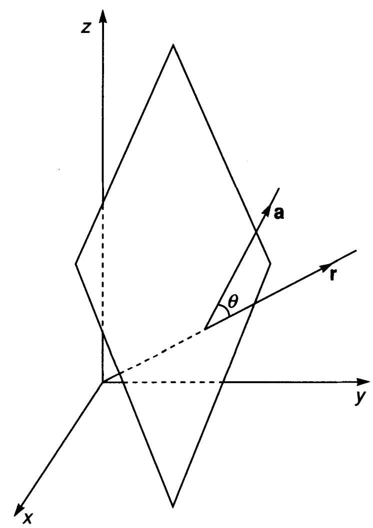


Figure 5.6 The plane $\mathbf{r} \cdot \mathbf{a} = c$.

which gives more geometrical insight than the Cartesian equation. The plane can be viewed as the collection of points that have the following common property: their projection on to the line joining the nearest point of this collection to the origin is the same number, $c/|\mathbf{a}|$. Secondly the equation $\mathbf{r} \cdot \mathbf{a} = c$ is completely co-ordinate free, and so does not depend on using Cartesian co-ordinates (for example). If we were to express **r** in another co-ordinate system (cylindrical polars for example, see Chapter 7), we could still use the equation $\mathbf{r} \cdot \mathbf{a} = c$ to find the equation of a plane. This is one of the most useful by-products of using vectors.

A sphere has the Cartesian equation $(x - b_1)^2 + (y - b_2)^2 + (z - b_3)^2 = k^2$. In terms of vectors **r** and $\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$ this is $|\mathbf{r} - \mathbf{b}|^2 = k^2$ or $|\mathbf{r} - \mathbf{b}| = k$.
Geometrically, this is the modulus of the vector $(\mathbf{r} - \mathbf{b})$ being constant (k), which is shown in Figure 5.7. Again the equation $|\mathbf{r} - \mathbf{b}| = k$ is co-ordinate free.

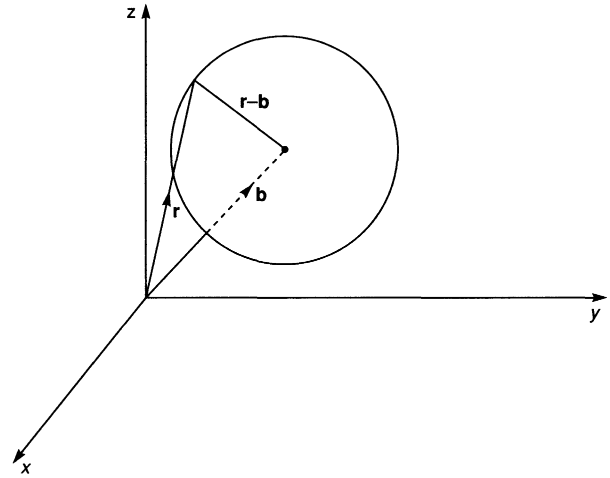


Figure 5.7 The sphere $|\mathbf{r} - \mathbf{b}|$ = constant.

Example 5.10 Show that the quadrilateral formed by joining the mid-points of the sides of an arbitrary quadrilateral must be a parallelogram.

Solution This quite remarkable result is very difficult to prove without the use of vectors, especially when it is realised that the arbitrary quadrilateral need not even be planar! Using vectors, however, the proof is quite straightforward and runs as follows: Let three sides of the arbitrary quadrilateral be described by the vectors **a**, **b**, and **c**. The fourth will then be **a** + **b** + **c** (see Figure 5.8).

The sides of the new quadrilateral formed by joining the mid-points of the sides are thus $\frac{1}{2}(\mathbf{a} + \mathbf{b})$, $\frac{1}{2}(\mathbf{b} + \mathbf{c})$, $\frac{1}{2}(\mathbf{a} + \mathbf{b} + \mathbf{c} - \mathbf{c})$ and $\frac{1}{2}(\mathbf{a} + \mathbf{b} + \mathbf{c} - \mathbf{a})$. From this, we see that two of the sides are represented by the vector $\frac{1}{2}(\mathbf{a} + \mathbf{b})$ and are thus parallel, and two are represented by the vector $\frac{1}{2}(\mathbf{b} + \mathbf{c})$ and are thus also parallel. This second quadrilateral is therefore, by definition, a parallelogram.

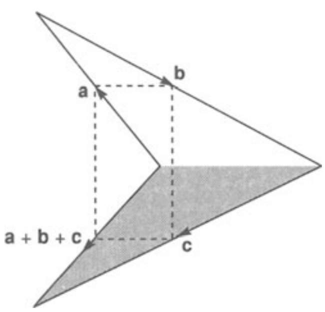


Figure 5.8 The 'dashed' quadrilateral is a parallelogram.

Example 5.11 Show that $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$.

Solution The definition of the cross product is $\mathbf{a} \times \mathbf{b} = |\mathbf{a}||\mathbf{b}| \sin\theta\, \hat{\mathbf{n}}$, where θ is the angle between the vectors **a** and **b**, and $\hat{\mathbf{n}}$ is a unit vector such that **a**, **b** and $\hat{\mathbf{n}}$ form a right-handed system (see Figure 5.9).

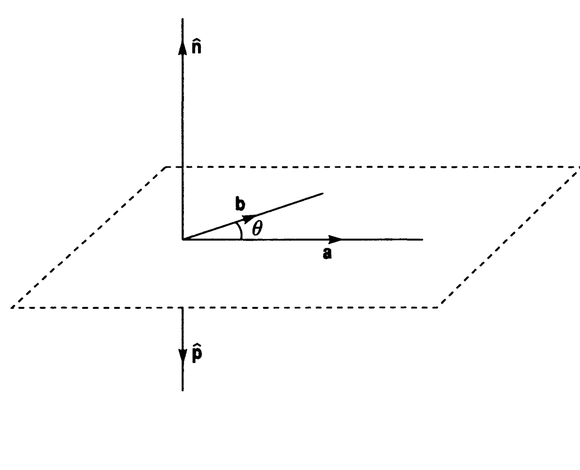


Figure 5.9 The right-handed systems **a**, **b**, $\hat{\mathbf{n}}$ and **b**, **a**, $\hat{\mathbf{p}}$

On the other hand, $\mathbf{b} \times \mathbf{a} = |\mathbf{b}||\mathbf{a}| \sin\theta\, \hat{\mathbf{p}}$ where the vectors $\mathbf{b}$, $\mathbf{a}$ and $\hat{\mathbf{p}}$ form a right-handed system. As both $\hat{\mathbf{n}}$ and $\hat{\mathbf{p}}$ are unit vectors perpendicular to the plane of $\mathbf{a}$ and $\mathbf{b}$ either $\hat{\mathbf{n}} = \hat{\mathbf{p}}$ or $\hat{\mathbf{n}} = -\hat{\mathbf{p}}$. In order that the triples $\mathbf{a}$, $\mathbf{b}$, $\hat{\mathbf{n}}$ and $\mathbf{b}$, $\mathbf{a}$, $\hat{\mathbf{p}}$ are both right-handed, we must have $\hat{\mathbf{n}} = -\hat{\mathbf{p}}$, thus

$$\mathbf{a} \times \mathbf{b} = |\mathbf{a}||\mathbf{b}| \sin\theta\, \hat{\mathbf{n}} = -|\mathbf{b}||\mathbf{a}| \sin\theta\, \hat{\mathbf{p}} = -\mathbf{b} \times \mathbf{a}$$

as required.

Example 5.12 Show that $\mathbf{a} \times \mathbf{a} = \mathbf{0}$.

Solution Since $\mathbf{a} \times \mathbf{a} = |\mathbf{a}||\mathbf{a}| \sin 0\, \hat{\mathbf{e}}$, the angle between $\mathbf{a}$ and itself being zero, and $\sin 0 = 0$, the result follows. In this example, $\hat{\mathbf{e}}$ is an arbitrary vector perpendicular to $\mathbf{a}$, but this information is not used.

Example 5.13 Establish formulae for all the vector products of the unit vectors $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ (dropping the carats).

Solution First of all, it is clear from the previous example that the vector (or cross) products of any of the vectors $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ with itself is zero. The vector product

$$\mathbf{i} \times \mathbf{j} = |\mathbf{i}||\mathbf{j}| \sin\left(\frac{\pi}{2}\right)\mathbf{k} = \mathbf{k} \text{ (for } \mathbf{i}, \mathbf{j} \text{ and } \mathbf{k} \text{ are a right-handed system)}$$

$$\mathbf{j} \times \mathbf{i} = -\mathbf{i} \times \mathbf{j} = -\mathbf{k} \text{ (using Example 5.11).}$$

Similarly, $\mathbf{j} \times \mathbf{k} = \mathbf{i}$ and $\mathbf{k} \times \mathbf{j} = -\mathbf{i}$, and also, $\mathbf{k} \times \mathbf{i} = \mathbf{j}$ and $\mathbf{i} \times \mathbf{k} = -\mathbf{j}$. The following table summarises all of these results:

$$\begin{aligned}
&\mathbf{i} \times \mathbf{i} = \mathbf{j} \times \mathbf{j} = \mathbf{k} \times \mathbf{k} = \mathbf{0}\\
&\mathbf{i} \times \mathbf{j} = \mathbf{k},\ \mathbf{j} \times \mathbf{k} = \mathbf{i},\ \mathbf{k} \times \mathbf{i} = \mathbf{j}\\
&\mathbf{j} \times \mathbf{i} = -\mathbf{k},\ \mathbf{k} \times \mathbf{j} = -\mathbf{i},\ \mathbf{i} \times \mathbf{k} = -\mathbf{j}
\end{aligned}$$

Example 5.14 Find an expression for $\mathbf{a} \times \mathbf{b}$ in terms of the components (a_1, a_2, a_3) and (b_1, b_2, b_3).

Solution This example is done using all the identities of Example 5.13. Writing $\mathbf{a}$ and $\mathbf{b}$ in component form, then calculating the cross product gives

$$\begin{aligned}
\mathbf{a} \times \mathbf{b} &= (a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}) \times (b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k})\\
&= a_1b_1\mathbf{i} \times \mathbf{i} + a_1b_2\mathbf{i} \times \mathbf{j} + a_1b_3\mathbf{i} \times \mathbf{k} +\\
&\quad a_2b_1\mathbf{j} \times \mathbf{i} + a_2b_2\mathbf{j} \times \mathbf{j} + a_2b_3\mathbf{j} \times \mathbf{k} +\\
&\quad a_3b_1\mathbf{k} \times \mathbf{i} + a_3b_2\mathbf{k} \times \mathbf{j} + a_3b_3\mathbf{k} \times \mathbf{k}\\
&= a_1b_2\mathbf{k} - a_1b_3\mathbf{j} - a_2b_1\mathbf{k} + a_2b_3\mathbf{i} + a_3b_1\mathbf{j} - a_3b_2\mathbf{i}\\
&\quad \text{using the previous example}
\end{aligned}$$

Rearranging this gives $\mathbf{a} \times \mathbf{b} = (a_2b_3 - a_3b_2)\mathbf{i} + (a_3b_1 - a_1b_3)\mathbf{j} + (a_1b_2 - a_2b_1)\mathbf{k}$ which can be written in the determinant form

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

Example 5.15 If $\mathbf{a}$ and $\mathbf{b}$ are fixed vectors that are not parallel, find all vectors $\mathbf{x}$ such that

$$\mathbf{a} \times \mathbf{x} = \mathbf{b} \times \mathbf{x}$$

Solution Given $\mathbf{a} \times \mathbf{x} = \mathbf{b} \times \mathbf{x}$, this implies $\mathbf{a} \times \mathbf{x} - \mathbf{b} \times \mathbf{x} = \mathbf{0}$ or

$$(\mathbf{a} - \mathbf{b}) \times \mathbf{x} = \mathbf{0}$$

This is only true if the vectors $\mathbf{a} - \mathbf{b}$ and $\mathbf{x}$ are parallel, that is $\mathbf{x} = k(\mathbf{a} - \mathbf{b})$. As a check

$$\mathbf{a} \times \mathbf{x} = \mathbf{a} \times k(\mathbf{a} - \mathbf{b}) = k\mathbf{a} \times \mathbf{a} - k\mathbf{a} \times \mathbf{b} = -k\mathbf{a} \times \mathbf{b}$$
$$\mathbf{b} \times \mathbf{x} = \mathbf{b} \times k(\mathbf{a} - \mathbf{b}) = k\mathbf{b} \times \mathbf{a} - k\mathbf{b} \times \mathbf{b} = k\mathbf{b} \times \mathbf{a}$$

and as $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$ these are indeed the same, thus $\mathbf{x} = k(\mathbf{a} - \mathbf{b})$ where k is an arbitrary scalar.

Example 5.16 Given $\mathbf{a} \times \mathbf{b} = \mathbf{a} \times \mathbf{c}$, does this imply that $\mathbf{b} = \mathbf{c}$?

Solution If $\mathbf{a} \times \mathbf{b} = \mathbf{a} \times \mathbf{c}$ then $\mathbf{a} \times (\mathbf{b} - \mathbf{c}) = \mathbf{0}$. This certainly does not mean that $\mathbf{b} = \mathbf{c}$, merely that $\mathbf{b} - \mathbf{c}$ is parallel to $\mathbf{a}$, or

$$\mathbf{b} = \mathbf{c} + k\mathbf{a}$$

where k is an arbitrary scalar.

Example 5.17 Find the solution to the equation $p\mathbf{x} + (\mathbf{x} \cdot \mathbf{b})\mathbf{a} = \mathbf{c}$ where p is a non-zero scalar, and $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ are constant vectors.

Solution If we take the dot (scalar) product of the equation with $\mathbf{b}$ we obtain

$$p\mathbf{x} \cdot \mathbf{b} + (\mathbf{x} \cdot \mathbf{b})(\mathbf{a} \cdot \mathbf{b}) = \mathbf{c} \cdot \mathbf{b}$$

Solving this for $\mathbf{x} \cdot \mathbf{b}$ gives

$$\mathbf{x} \cdot \mathbf{b} = \frac{\mathbf{c} \cdot \mathbf{b}}{p + (\mathbf{a} \cdot \mathbf{b})}, \text{ provided } p \neq -(\mathbf{a} \cdot \mathbf{b})$$

From the original equation, we get

$$\mathbf{x} = \frac{\mathbf{c}}{p} - (\mathbf{x} \cdot \mathbf{b}) \frac{\mathbf{a}}{p}, \text{ provided } p \neq 0$$

Substituting for $\mathbf{x} \cdot \mathbf{b}$ yields

$$\mathbf{x} = \frac{\mathbf{c}}{p} - \frac{(\mathbf{c} \cdot \mathbf{b})\mathbf{a}}{p^2 + p(\mathbf{a} \cdot \mathbf{b})}$$

As is typical in mathematics, we spend more time tying up the loose ends than deriving the above solution (which is valid almost always)! First of all, the case $p = -(\mathbf{a} \cdot \mathbf{b})$ is an interesting red herring. With p equal to this value, the left-hand side of the original equation is $(\mathbf{x} \cdot \mathbf{b})\mathbf{a} - (\mathbf{a} \cdot \mathbf{b})\mathbf{x}$ which is in fact the expansion of a vector triple product $\mathbf{b} \times (\mathbf{a} \times \mathbf{x})$ (see Example 5.22). The problem therefore becomes one of solving the equation $\mathbf{b} \times (\mathbf{a} \times \mathbf{x}) = \mathbf{c}$. From this equation, $\mathbf{c}$ is perpendicular to both $\mathbf{b}$ and $(\mathbf{a} \times \mathbf{x})$. However, $(\mathbf{a} \times \mathbf{x})$ is itself perpendicular to both $\mathbf{a}$ and $\mathbf{x}$. Therefore we must conclude that the three vectors $\mathbf{a}$, $\mathbf{x}$ and $\mathbf{c}$ lie in a plane that is perpendicular to $\mathbf{b}$. Hence $\mathbf{a} \cdot \mathbf{b} = 0$, and so $p = -(\mathbf{a} \cdot \mathbf{b})$ is also zero. Moreover, $\mathbf{x} \cdot \mathbf{b}$ must be zero ($\mathbf{x}$ is perpendicular to $\mathbf{b}$) and so $\mathbf{x}$ has disappeared completely from the original problem which now implies $\mathbf{c} = \mathbf{0}$. (If $\mathbf{c} = \mathbf{0}$, then $\mathbf{b} \times (\mathbf{a} \times \mathbf{x}) = \mathbf{0}$ and $\mathbf{x}$ is any vector in the plane of $\mathbf{a}$.)

Next we look at the case $p = 0$, which implies

$$(\mathbf{x} \cdot \mathbf{b})\mathbf{a} = \mathbf{c}$$

which is equivalent to $\mathbf{x} \cdot \mathbf{b} = \lambda$ (see the next example), $\mathbf{a}$ being parallel to $\mathbf{c}$.

Example 5.18 Find the general solution of the equation $\mathbf{x} \cdot \mathbf{a} = \lambda$ where λ is a scalar and $\mathbf{a}$ a fixed vector.

Solution Since $\mathbf{a}$ is an arbitrary but constant vector, the x-axis can, without loss of generality, be aligned in its direction so that $\mathbf{a} = a\hat{\mathbf{i}}$. In general, $\mathbf{x} = x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}$ in component form, therefore $\mathbf{x} \cdot \mathbf{a} = x_1 a$, so $x_1 = \dfrac{\lambda}{a}$ and

$$\mathbf{x} = \frac{\lambda\mathbf{i}}{a} + x_2\mathbf{k} \times \mathbf{i} - x_3\mathbf{j} \times \mathbf{i}$$

$$\mathbf{x} = \frac{\lambda\mathbf{a}}{a^2} + \mathbf{h} \times \mathbf{i} = \frac{\lambda\mathbf{a}}{a^2} + \frac{\mathbf{h} \times \mathbf{a}}{a} \quad \text{writing } \mathbf{h} = x_2\hat{\mathbf{k}} - x_3\hat{\mathbf{j}}$$

$$= \mathbf{c} \times \mathbf{a} + \frac{\lambda\mathbf{a}}{a^2}$$

where $\mathbf{c}$ is an arbitrary vector. (It can be in any direction, but only its components that are perpendicular to $\mathbf{a}$ feature in the solution since the cross product $\mathbf{c} \times \mathbf{a}$ eliminates that in the $\mathbf{a}$ direction.)

Example 5.19 The scalar triple product is defined as $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$. Show that

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{b} \cdot (\mathbf{c} \times \mathbf{a}) = \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b})$$

Solution Since the vector product $\mathbf{b} \times \mathbf{c}$ can be calculated by evaluating the determinant

$$\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} \quad \text{(see Example 5.13)}$$

which can be expanded as $\mathbf{i}\begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} + \mathbf{j}\begin{vmatrix} b_3 & b_1 \\ c_3 & c_1 \end{vmatrix} + \mathbf{k}\begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}$, the scalar product of this with $\mathbf{a}$ gives

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = a_1\begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} + a_2\begin{vmatrix} b_3 & b_1 \\ c_3 & c_1 \end{vmatrix} + a_3\begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}$$

$$= \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

The other scalar products are thus

$$\mathbf{b} \cdot (\mathbf{c} \times \mathbf{a}) = \begin{vmatrix} b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \\ a_1 & a_2 & a_3 \end{vmatrix} \text{ and } \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}) = \begin{vmatrix} c_1 & c_2 & c_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

Since the value of any determinant is the same if its rows are permuted cyclically, the three scalar products $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$, $\mathbf{b} \cdot (\mathbf{c} \times \mathbf{a})$ and $\mathbf{c} \cdot (\mathbf{a} \times \mathbf{b})$ are all equal.

Also, since $\mathbf{b} \times \mathbf{c} = -\mathbf{c} \times \mathbf{b}$, we have $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = -\mathbf{a} \cdot (\mathbf{c} \times \mathbf{b})$ and using the same determinant arguments as above we can deduce that

$$\mathbf{a} \cdot (\mathbf{c} \times \mathbf{b}) = \mathbf{c} \cdot (\mathbf{b} \times \mathbf{a}) = \mathbf{b} \cdot (\mathbf{a} \times \mathbf{c})$$

Example 5.20 Find the volume of the parallelepiped formed by three arbitrary vectors $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ whose ends coincide at the origin and which are not coplanar.

Solution The parallelepiped (a solid all of whose faces are parallelograms) is shown in Figure 5.10.

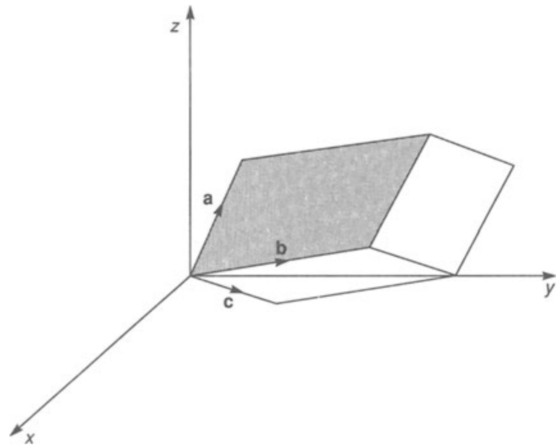


Figure 5.10 A parallelepiped with sides **a**, **b**, **c**.

Since $\mathbf{a} \times \mathbf{b} = |\mathbf{a}||\mathbf{b}| \sin\theta\, \hat{\mathbf{n}}$, where θ is the angle between **a** and **b**, the area of the shaded parallelogram in Figure 5.10 must be $|\mathbf{a} \times \mathbf{b}|$ (the area of a parallelogram is the product of the lengths of adjacent sides multiplied by the sine of the angle between them). If the 'height' of the parallelepiped is denoted by h then its volume is $h|\mathbf{a} \times \mathbf{b}|$.

Now, $\mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}) = |\mathbf{c}||\mathbf{a} \times \mathbf{b}| \cos\phi$ where ϕ is the angle between **c** and the (unit) normal to the plane of **a** and **b**. However, $h = |\mathbf{c}| \cos\phi$, hence the volume of the parallelepiped is $(|\mathbf{c}| \cos\phi)\, |\mathbf{a} \times \mathbf{b}| = \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b})$. This gives a very convenient geometrical interpretation of the scalar triple product. In particular, it is now obvious why the scalar triple products $\mathbf{c} \cdot (\mathbf{a} \times \mathbf{b})$, $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$ and $\mathbf{b} \cdot (\mathbf{c} \times \mathbf{a})$ must all be equal, since they represent the same volume. Less obvious is why this volume becomes negative when the order of two of the vectors in any triple is changed, but it can be traced back to the interpretation of the direction of the normal in the definition of the cross product. For most applications, the negative sign is simply ignored in the calculation of volumes. Another by-product is the result that three vectors **a**, **b** and **c** are coplanar if and only if the scalar triple product $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$ is zero.

Example 5.21 Use vector algebra to determine the shortest distance between two skew lines.

Solution In order to solve this particular problem, we need a to revise a little background geometry. It does not matter if it is not actually revision, we hope it is reasonably easy to follow, and more about using vectors to describe geometry will in any case be found in the next chapter. A straight line can be represented by the expression $\mathbf{r} = \mathbf{a} + \mathbf{b}t$ where **r** is the position vector and **a** and **b** are (non-parallel) constant vectors, and t is a scalar variable. Figure 5.11 shows this straight line schematically.

The position vector **r** is, as its name implies, the vector of the position of an arbitrary point on the line, and, as t varies, this point runs the length of the line. The value $t = 0$ corresponds to the

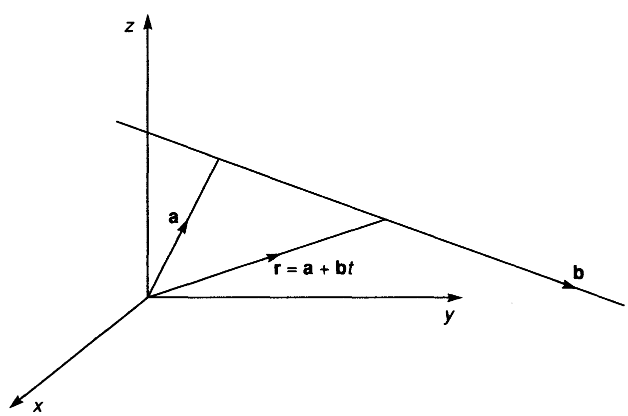


Figure 5.11 The straight line $\mathbf{r} = \mathbf{a} + \mathbf{b}t$.

vector $\mathbf{a}$ and $t = \infty$ corresponds to the vector $\mathbf{b}$ which can be seen to be the direction of the line itself. However, it is unwise to refer to the straight line as '$\mathbf{b}$' as $\mathbf{b}$ is not line bound. The question demands the use of line bound vectors through its requirement to find a length. We need to postulate a second line, let this be given by the position vector $\mathbf{r} = \mathbf{a}' + \mathbf{b}'s$ where $\mathbf{a}'$ and $\mathbf{b}'$ are two more non-parallel vectors and s is another scalar variable independent of t. Additionally, neither $\mathbf{a}$ or $\mathbf{a}'$ nor $\mathbf{b}$ or $\mathbf{b}'$ are parallel. For convenience, call the first line l and the second line l'. Of course, both l and l' are line bound once $\mathbf{a}$, $\mathbf{a}'$, $\mathbf{b}$ and $\mathbf{b}'$ are fixed since $\mathbf{r}$ the position vector has one end at the origin.

The vector $\mathbf{b} \times \mathbf{b}'$ is perpendicular to both l and l'. Let A be the point with position vector $\mathbf{r} = \mathbf{a}$ and A' be the point with position vector $\mathbf{r} = \mathbf{a}'$. The vector $\overrightarrow{AA}$ is then $\mathbf{a}' - \mathbf{a}$. The shortest distance between l and l' is along a line that is perpendicular to both l and l'. Therefore this line of shortest distance is parallel to $\mathbf{b} \times \mathbf{b}'$. If $\hat{\mathbf{n}}$ is the unit vector in this direction, then

$$\hat{\mathbf{n}} = \frac{\mathbf{b} \times \mathbf{b}'}{|\mathbf{b} \times \mathbf{b}'|}$$

and the length of the perpendicular line is the required shortest distance. This must be the projection of $\overrightarrow{AA}$ on to $\hat{\mathbf{n}}$, that is $(\mathbf{a}' - \mathbf{a}) \cdot \hat{\mathbf{n}}$, or

$$(\mathbf{a}' - \mathbf{a}) \cdot \frac{\mathbf{b} \times \mathbf{b}'}{|\mathbf{b} \times \mathbf{b}'|}$$

Two points are worth emphasising here. First of all, *any* point on l connecting *any* point on l' (we chose AA') would do to project on to the vector $\hat{\mathbf{n}}$. Secondly, the above distance could turn out to be negative; this depends on the direction of $\mathbf{b} \times \mathbf{b}'$ and $\mathbf{a}' - \mathbf{a}$. The modulus gives the correct answer in all cases.

Example 5.22 Find the shortest distance between the two lines $\dfrac{x-1}{2} = \dfrac{y}{3} = \dfrac{z+1}{1}$, and $\dfrac{x}{3} = \dfrac{y-3}{4} = \dfrac{z-1}{2}$.

Solution To find the shortest distance, we utilise the formula derived in the previous example. In order to do this we need to put the equations of the two lines into the form $\mathbf{r} = \mathbf{a} + \mathbf{b}t$. (This is in fact the principal reason for this example.)

The straight line $\dfrac{x-1}{2} = \dfrac{y}{3} = \dfrac{z+1}{1}$ can be written as three equations (*not* independent):

$$3x - 2y - 3 = 0$$
$$y - 3z - 3 = 0$$
$$x - 2z - 3 = 0$$

Letting $y = 3t$ which is equivalent to putting $\dfrac{x-1}{2} = \dfrac{y}{3} = \dfrac{z+1}{1} = t$ implies $x = 2t + 1$ and $z = t - 1$. (In fact, putting $x = t$ then getting y in terms of t then z in terms of t works just as well, but $y = 3t$ avoids fractions.) Writing these as

$$x = 1 + 2t$$
$$y = 0 + 3t$$
$$z = -1 + t$$

gives $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k} = (1, 0, -1) + (2, 3, 1)t$. Thus the line is in the required form with $\mathbf{a} = (1, 0, -1)$ and $\mathbf{b} = (2, 3, 1)$.

Proceeding similarly with the second line $\dfrac{x}{3} = \dfrac{y-3}{4} = \dfrac{z-1}{2}$ leads to

$$4x - 3y + 9 = 0$$
$$2x - 3z + 3 = 0$$
$$y - 2z - 1 = 0$$

This time, putting $x = 3s$ is convenient and gives $y = 4s + 3$ and $z = 2s + 1$ so

$$x = 0 + 3s$$
$$y = 3 + 4s$$
$$z = 1 + 2s$$

giving $\mathbf{r} = (0, 3, 1) + (3, 4, 2)s$ so that $\mathbf{a}' = (0, 3, 1)$ and $\mathbf{b}' = (3, 4, 2)$. The perpendicular normal is given by $\hat{\mathbf{n}} = \dfrac{\mathbf{b} \times \mathbf{b}'}{|\mathbf{b} \times \mathbf{b}'|}$ which is calculated as follows:

$$\mathbf{b} \times \mathbf{b}' = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 2 & 3 & 1 \\ 3 & 4 & 2 \end{vmatrix} = (2, -1, -1)$$

$$\text{so } \hat{\mathbf{n}} = \frac{1}{\sqrt{6}} (2, -1, -1)$$

The vector $\mathbf{a}' - \mathbf{a} = (0, 3, 1) - (1, 0, -1) = (-1, 3, 2)$. The shortest distance is therefore p where

$$p = (\mathbf{a}' - \mathbf{a}) \cdot \hat{\mathbf{n}} = \frac{1}{\sqrt{6}} (-2 -3 -2) = -\frac{7}{\sqrt{6}}$$

We ignore the negative sign as p is a length (the reasons are elaborated at the end of the last example). Hence the required perpendicular distance is $\dfrac{7}{\sqrt{6}} = 2.85$.

Example 5.23 Find the shortest distance between two opposite sides of a regular tetrahedron of side l.

Solution Let the tetrahedron be as shown in Figure 5.12. Let us select $\mathbf{b}$ and $\mathbf{f}$ as the two sides that we wish to find the shortest distance between. Now $\mathbf{b} \times \mathbf{f}$ is perpendicular to both $\mathbf{b}$ and $\mathbf{f}$, and since $\mathbf{a}$ links both, the desired shortest distance is $p = l\hat{\mathbf{a}}(\hat{\mathbf{b}} \times \hat{\mathbf{f}})$. In order to evaluate this, first note that all lengths of sides are equal to l. Second, that $\hat{\mathbf{a}} \cdot (\hat{\mathbf{b}} \times \hat{\mathbf{f}}) = \hat{\mathbf{a}} \cdot (\hat{\mathbf{b}} \times (\hat{\mathbf{c}} - \hat{\mathbf{a}})) = \hat{\mathbf{a}} \cdot (\hat{\mathbf{b}} \times \hat{\mathbf{c}})$. If $\hat{\mathbf{n}}$ is a unit normal to the plane of $\mathbf{b}$ and $\mathbf{c}$, the required distance is

$$p = l\hat{\mathbf{a}} \cdot \hat{\mathbf{n}} \sin\frac{\pi}{3} = \frac{l\sqrt{3}}{2}\, \hat{\mathbf{a}} \cdot \hat{\mathbf{n}} = \frac{l\sqrt{3}}{2} \cos\theta$$

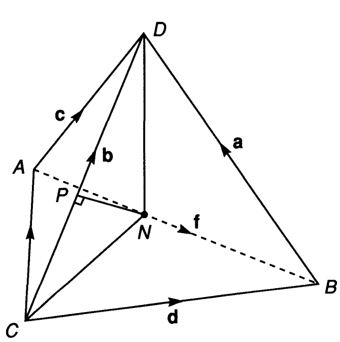


Figure 5.12 A regular tetrahedron.

where θ is the angle between **a** and **n**. This angle is a little tricky to find, in fact it is easier to find $\frac{\pi}{2} - \theta$. In Figure 5.13, we have drawn triangle CDN where N is the mid-point of the side AB. If P is the foot of the perpendicular from N to CD then we can find the desired angle, shown in the figure as $D\hat{N}P$, as follows:

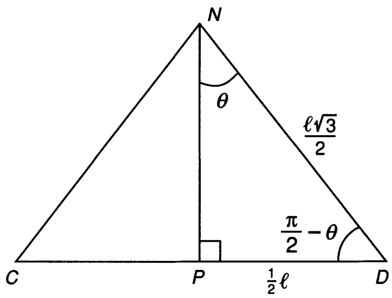


Figure 5.13 The triangle CND and the angle θ.

Since $PD = \frac{1}{2}l$ we have $\sin\theta = \dfrac{1}{\sqrt{3}}$ and $\cos\theta = \dfrac{\sqrt{2}}{\sqrt{3}}$. Therefore

$$p = \frac{l\sqrt{3}}{2}\frac{\sqrt{2}}{\sqrt{3}} = \frac{l}{\sqrt{2}}$$

The length of the shortest distance is therefore $\dfrac{l}{\sqrt{2}}$.

Example 5.24 Prove that $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) \equiv (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}$ for any three vectors **a**, **b** and **c** where the three lines indicate identity.

Solution From its definition, the cross product $\mathbf{b} \times \mathbf{c}$ is a vector that is perpendicular to the two vectors **b** and **c**. The cross product $\mathbf{a} \times (\mathbf{b} \times \mathbf{c})$ is similarly a vector that is perpendicular to the two vectors **a** and $\mathbf{b} \times \mathbf{c}$. These two statements imply that $\mathbf{a} \times (\mathbf{b} \times \mathbf{c})$ is in the plane of **b** and **c** (both are perpendicular to $\mathbf{b} \times \mathbf{c}$). Hence we can write

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) \equiv l\mathbf{b} + m\mathbf{c}$$

where l and m are (scalar) constants to be determined. Take the scalar product of this equation with **a** to obtain

$$0 = l(\mathbf{a} \cdot \mathbf{b}) + m(\mathbf{a} \cdot \mathbf{c})$$

the left-hand side being zero since $\mathbf{a} \times (\mathbf{b} \times \mathbf{c})$ is perpendicular to **a**. Thus

$$m = -\frac{l(\mathbf{a} \cdot \mathbf{b})}{(\mathbf{a} \cdot \mathbf{c})}$$

or, writing $\mu = \dfrac{l}{(\mathbf{a} \cdot \mathbf{c})}$ for another constant

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) \equiv \mu[(\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}]$$

The constant μ can only be determined by substituting specific values for **a**, **b** and **c** (those familiar with dimensional analysis will have met an analogous situation). Letting $\mathbf{a} = \mathbf{b} = \mathbf{i}$, $\mathbf{c} = \mathbf{j}$ gives $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \mathbf{i} \times (\mathbf{i} \times \mathbf{j}) = \mathbf{i} \times \mathbf{k} = -\mathbf{j}$. Also, $\mu[(\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}] = \mu[(\mathbf{i} \cdot \mathbf{j})\mathbf{i} - (\mathbf{i} \cdot \mathbf{i})\mathbf{j}] = -\mu\mathbf{j}$. Hence $\mu = 1$ and the result is established.

Example 5.25 Show that any vector $\mathbf{v}$ can be expressed in the form

$$\mathbf{v} = \frac{[(\mathbf{v} \times \mathbf{b}) \cdot \mathbf{c}]\mathbf{a} + [(\mathbf{v} \times \mathbf{c}) \cdot \mathbf{a}]\mathbf{b} + [(\mathbf{v} \times \mathbf{a}) \cdot \mathbf{b}]\mathbf{c}}{(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}},$$

where $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ are arbitrary vectors.

Solution This problem is solved by looking at the vector product of four vectors as follows:

$$(\mathbf{a} \times \mathbf{b}) \times (\mathbf{c} \times \mathbf{v}) = (\mathbf{a} \times \mathbf{b}) \times \mathbf{p} \text{ where } \mathbf{p} = \mathbf{c} \times \mathbf{v}$$
$$= (\mathbf{a} \cdot \mathbf{p})\mathbf{b} - (\mathbf{b} \cdot \mathbf{p})\mathbf{a} \text{ using the previous result.}$$

Alternatively we expand the four vector as the different triple product $\mathbf{s} \times (\mathbf{c} \times \mathbf{v})$ where $\mathbf{s} = \mathbf{a} \times \mathbf{b}$ to give

$$(\mathbf{a} \times \mathbf{b}) \times (\mathbf{c} \times \mathbf{v}) = \mathbf{s} \times (\mathbf{c} \times \mathbf{v})$$
$$= (\mathbf{s} \cdot \mathbf{v})\mathbf{c} - (\mathbf{s} \cdot \mathbf{c})\mathbf{v}$$

These two right-hand sides are now equated

$$(\mathbf{a} \cdot \mathbf{p})\mathbf{b} - (\mathbf{b} \cdot \mathbf{p})\mathbf{a} = (\mathbf{s} \cdot \mathbf{v})\mathbf{c} - (\mathbf{s} \cdot \mathbf{c})\mathbf{v}$$

or

$$(\mathbf{a} \cdot \mathbf{c} \times \mathbf{v})\mathbf{b} - (\mathbf{b} \cdot \mathbf{c} \times \mathbf{v})\mathbf{a} = (\mathbf{a} \times \mathbf{b} \cdot \mathbf{v})\mathbf{c} - (\mathbf{a} \times \mathbf{b} \cdot \mathbf{c})\mathbf{v}$$

Notice the absence of parentheses in the scalar triple products, this is because it is impossible to evaluate the scalar triple product in more than one way (you cannot take the vector product of a scalar!). Rearranging to make $\mathbf{v}$ the subject of this last equation gives

$$(\mathbf{a} \times \mathbf{b} \cdot \mathbf{c})\mathbf{v} = (\mathbf{a} \times \mathbf{b} \cdot \mathbf{v})\mathbf{c} + (\mathbf{b} \cdot \mathbf{c} \times \mathbf{v})\mathbf{a} - (\mathbf{a} \cdot \mathbf{c} \times \mathbf{v})\mathbf{b}$$

or, using the properties of the scalar triple product

$$(\mathbf{a} \times \mathbf{b} \cdot \mathbf{c})\mathbf{v} = (\mathbf{v} \times \mathbf{b} \cdot \mathbf{c})\mathbf{a} + (\mathbf{v} \times \mathbf{c} \cdot \mathbf{a})\mathbf{b} + (\mathbf{v} \times \mathbf{a} \cdot \mathbf{b})\mathbf{c}$$

from which, on division by $(\mathbf{a} \times \mathbf{b} \cdot \mathbf{c})$, a scalar of course, the result follows.

5.3 Exercises

5.1. Find the sum and difference of the following pairs of vectors:
(a) $\mathbf{i} + 5\mathbf{j}$, $5\mathbf{i} + \mathbf{j}$,
(b) $\mathbf{i} - \mathbf{j}$, $\mathbf{i} + \mathbf{j}$,
(c) $\mathbf{i} - \mathbf{j} - \mathbf{k}$, $\mathbf{i} + \mathbf{j} - \mathbf{k}$,
(d) $2\mathbf{i} + \mathbf{j} + 2\mathbf{k}$, $-3\mathbf{i} + 3\mathbf{j} - 3\mathbf{k}$,
(e) $\mathbf{i} + \mathbf{k}$, $\mathbf{j} - \mathbf{k}$.
Find also the *unit* vectors that correspond to each of the ten vectors given above.

5.2. Determine the direction cosines of all the vectors in Exercise 5.1.

5.3. If $\mathbf{a} = 2\mathbf{i} - \mathbf{j} + \mathbf{k}$, $\mathbf{b} = \mathbf{i} + 3\mathbf{j} - 2\mathbf{k}$, $\mathbf{c} = -2\mathbf{i} + \mathbf{j} - 3\mathbf{k}$ and $\mathbf{d} = 3\mathbf{i} + 2\mathbf{j} + 5\mathbf{k}$ determine constants α, β and γ such that $\mathbf{d} = \alpha\mathbf{a} + \beta\mathbf{b} + \gamma\mathbf{c}$. This is called *linear dependence*. (Four vectors are always linearly dependent because we live in three dimensions.)

5.4. Show that the vectors $\mathbf{a} = \mathbf{i} + 2\mathbf{j} + 3\mathbf{k}$, $\mathbf{b} = 3\mathbf{i} - 3\mathbf{j} + \mathbf{k}$, $\mathbf{c} = 2\mathbf{i} - 5\mathbf{j} - 2\mathbf{k}$ form the sides of a right-angled triangle, and find the other two angles.

5.5. Determine the two values of θ ($0 \leq \theta \leq 2\pi$) for which the lines $\mathbf{i}\cos\theta - \mathbf{j}\sin\theta + \mathbf{k}$ and $\mathbf{i}\cos\theta + \mathbf{j}\sin\theta + \mathbf{k}$ are perpendicular.

5.6. The vectors $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ are given by

$$\mathbf{a} = (1, 0, -1),\ \mathbf{b} = (1, 1, 0) \text{ and } \mathbf{c} = (0, 1, 2)$$

Determine the values of

$$\mathbf{a} \cdot (\mathbf{b} + \mathbf{c});\ (\mathbf{a} - \mathbf{b}) \cdot \mathbf{c};\ (\mathbf{a} + \mathbf{b} + \mathbf{c}) \cdot (\mathbf{a} - \mathbf{b} + \mathbf{c})$$

5.7. Find the angles between the five pairs of vectors given in Exercise 5.1.

5.8. Determine the value of the constant p such that the two vectors $\mathbf{a} = p\mathbf{i} + \mathbf{j} - \mathbf{k}$ and $\mathbf{b} = \mathbf{i} + p\mathbf{j} - \mathbf{k}$ are
(a) at right angles,
(b) at an angle $\frac{\pi}{3}$,
to each other.

5.9. Find a vector that is perpendicular to the plane of the two vectors $\mathbf{i} - \mathbf{j} + \mathbf{k}$ and $\mathbf{i} + \mathbf{j} - \mathbf{k}$ and has magnitude $\sqrt{2}$.

5.10. Represent the two unit vectors $\mathbf{r}_1 = \cos\alpha\mathbf{i} - \sin\alpha\mathbf{j}$, $\mathbf{r}_2 = \cos\beta\mathbf{i} + \sin\beta\mathbf{j}$ in the x–y plane, and by considering their scalar and vector products establish the trigonometric formulae $\cos(\alpha + \beta) = \cos\alpha \cos\beta - \sin\alpha \sin\beta$ and $\sin(\alpha + \beta) = \sin\alpha \cos\beta + \cos\alpha \sin\beta$.

5.11. If $\mathbf{a} = 2\mathbf{i} - \mathbf{j} + \mathbf{k}$ and $\mathbf{b} = \mathbf{i} - 3\mathbf{j} - \mathbf{k}$ determine the vectors $\mathbf{a} \times \mathbf{b}$; $\mathbf{b} \times \mathbf{a}$; $(\mathbf{a} - \mathbf{b}) \times (\mathbf{a} + \mathbf{b})$. Establish in general that $2\mathbf{a} \times \mathbf{b} = (\mathbf{a} - \mathbf{b}) \times (\mathbf{a} + \mathbf{b})$.

5.12. The vector $\mathbf{F} = (3, 2, 4)$ represents a force. Find:
(a) its component in the $\mathbf{i}$ direction,
(b) its component in the direction $(1, 2, 3)$,
(c) the work done by the force $\mathbf{F}$ in moving unit mass along the line connecting $P(0, 0, 1)$ to $Q(2, 5, 7)$.
(In mechanics, the component of $\mathbf{F}$ in a direction making an angle θ to the direction of $\mathbf{F}$ is $|\mathbf{F}| \cos\theta$. Work done is force times distance, which in terms of the vector algebra of this chapter is the scalar product of $\mathbf{F}$ with the vector that represents the (straight) path of the mass. For a more general definition of work done we wait until Chapter 8.)

5.13. Prove that the vector sum of the three diagonals of the faces of a cube that spring from one vertex is equal to twice the vector that represents the diagonal of the cube itself that emerges from the same corner.

5.14. Determine the vector equation of the straight line connecting the points $(1, -2, 1)$ and $(0, -2, 3)$. Determine also the equation of the plane containing the origin and the points $(2, 4, 1)$ and $(4, 0, 2)$. Hence find the point where the line intersects the plane.

5.15. If P, Q, R are three points in space, prove that the area of the triangle PQR is $\frac{1}{2}\left|\overrightarrow{PQ} \times \overrightarrow{QR}\right|$. Hence determine the area of the triangle connecting the points:
(a) $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$,
(b) $(1, 2, 1)$, $(3, 0, 0)$, $(2, 1, 3)$.

5.16. A triangle has vertices A $(1, 2, 4)$, B $(-2, 2, 1)$ and C $(2, 4, -3)$. Show that $\cos A = \frac{1}{3}\sqrt{3}$, $\cos C = \frac{1}{3}\sqrt{6}$ and B is a right angle.

5.17. Find the unit vector $\hat{\mathbf{n}}$ that is perpendicular to both of the vectors $\mathbf{i} - \mathbf{j} - 2\mathbf{k}$ and $2\mathbf{i} + 2\mathbf{j} - \mathbf{k}$.

5.18. Show that an alternative expression for the vector area of a triangle whose vertices are at the points $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ is $\frac{1}{2}(\mathbf{b} \times \mathbf{c} + \mathbf{c} \times \mathbf{a} + \mathbf{a} \times \mathbf{b})$. Interpret the condition imposed on these points by the vanishing of this expression.

5.19. If $\hat{\mathbf{n}}$ is the unit vector along the bisector of the angle between the two unit vectors $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$, prove that $2(\hat{\mathbf{n}} \cdot \hat{\mathbf{a}})\hat{\mathbf{n}} = \hat{\mathbf{a}} + \hat{\mathbf{b}}$.

5.20. For arbitrary vectors $\mathbf{a}$, $\mathbf{b}$ show that

$$(\mathbf{a} \times \mathbf{b})^2 = \mathbf{a}^2\mathbf{b}^2 - (\mathbf{a} \cdot \mathbf{b})^2$$

5.21. Show that the general solution of the equation

$$a\mathbf{x}^2 + \mathbf{b} \cdot \mathbf{x} + c = 0$$

is given by $$\mathbf{x} = -\frac{\mathbf{b}}{2a} + \frac{\hat{\mathbf{e}}}{2a}\sqrt{\mathbf{b}^2 - 4ac}$$
where $\hat{\mathbf{e}}$ is an arbitrary unit vector.

5.22. Show that the line joining the points A $(2, -3, -1)$ and B $(8, -1, 2)$ has equations:

$$\tfrac{1}{6}(x - 2) = \tfrac{1}{2}(y + 3) = \tfrac{1}{3}(z + 1)$$

Find two points on this line whose distance from A is 14.

5.23. Find the distance from the point $(1, 4, -2)$ to the straight line through the point $(3, 1, -4)$ and parallel to the vector $(6, -2, 3)$.

5.24. Show that the plane through the point $(2, -4, 5)$ perpendicular to the line of intersection of the planes $2x + 3y - 4z = 1$ and $3x + y - 2z = 2$ is $2x + 8y + 7z = 7$.

5.25. Find the equation of the line of intersection of the planes $3x - y + z = 12$ and $x + 4y - 2z + 5 = 0$. Also find the plane which is perpendicular to this line and contains the point $(1, 2, -1)$.

5.26. Find the perpendicular distance between the diagonals AE and FG of the cube shown in Figure 5.14.

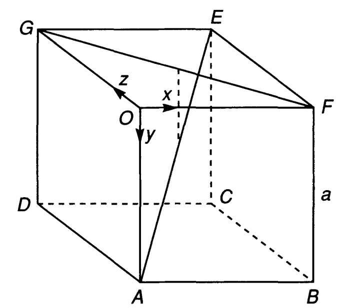


Figure 5.14 A cube of side a.

Find also the equations of these two diagonals.

5.27. Find the centre of the sphere inscribed in the tetrahedron bounded by the planes $x = 0$, $y = 0$, $z = 0$, $x + y + z = a$, where a is a constant.

5.28. Using vector algebra, from the equations

$$\mathbf{r}' = \mathbf{r} + \left[\frac{\gamma - 1}{\mathbf{v}^2}\,\mathbf{v} \cdot \mathbf{r} - \gamma t\right]\mathbf{v}$$

$$t' = \gamma\left(t - \frac{\mathbf{v} \cdot \mathbf{r}}{c^2}\right)$$

obtain $\mathbf{r}$ and t in terms of $\mathbf{r}'$ and t'. [These are transformation equations from Einstein's Special Theory of Relativity; $\gamma = \dfrac{c}{\sqrt{c^2 - \mathbf{v}^2}}$ and c is the speed of light.]

5.29. Show that the equation of a sphere that has points $\mathbf{a}$ and $\mathbf{b}$ as extremities of a diameter is $(\mathbf{r} - \mathbf{a}) \cdot (\mathbf{r} - \mathbf{b}) = 0$.

5.30. Using a method similar to that of Examples 5.21 and 5.22, find the shortest distance between the two skew lines

$$\frac{x + 3}{4} = \frac{y - 3}{-1} = \frac{z - 2}{1}$$

and $$\frac{x - 1}{2} = \frac{y - 5}{1} = \frac{z + 3}{2}$$

Topic Guide

Fields
Rules for Differentiation
Differential Geometry
Serret–Frenet Formulae
Mechanics

6 Vector Differentiation

6.1 Fact Sheet

A vector **F** that depends on a variable t is called a *vector function* (of a scalar variable – there are vector functions of vector variables, but not yet) and is written $\mathbf{F}(t)$. The derivative of $\mathbf{F}(t)$ (with respect to t of course) is written $\dfrac{\mathrm{d}\mathbf{F}}{\mathrm{d}t}$ and is defined by $\lim\limits_{\Delta t\to 0}\left\{\dfrac{\mathbf{F}(t+\Delta t)-\mathbf{F}(t)}{\Delta t}\right\}$ provided the limit exists and is unique. The formal definition of a column matrix made in Chapter 4 can be used to define a vector, and in that chapter we extended the notion to functions. The form of vectors introduced there are those found in books on linear algebra; by contrast, in this chapter we shall concentrate on three-dimensional vectors that are used to represent geometrical objects and their movement (that is, mechanics) in three-dimensional space. The vector functions defined here and the column matrices defined in Chapter 4 (provided they have three rows) are the same mathematically, but it is unusual and in my view confusing to couch three-dimensional applications of vectors in terms of matrix algebra. The notation used in this chapter is standard throughout engineering and applied science.

The scalar and vector products of two vector functions $\mathbf{F}(t)$ and $\mathbf{G}(t)$ differentiate in a similar way to the product of two scalar functions, namely

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{F}(t)\cdot\mathbf{G}(t)) = \frac{\mathrm{d}\mathbf{F}(t)}{\mathrm{d}t}\cdot\mathbf{G}(t) + \mathbf{F}(t)\cdot\frac{\mathrm{d}\mathbf{G}(t)}{\mathrm{d}t}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{F}(t)\times\mathbf{G}(t)) = \frac{\mathrm{d}\mathbf{F}(t)}{\mathrm{d}t}\times\mathbf{G}(t) + \mathbf{F}(t)\times\frac{\mathrm{d}\mathbf{G}(t)}{\mathrm{d}t}$$

(Be careful to retain the order of cross products as they do not commute $\mathbf{a}\times\mathbf{b} = -\mathbf{b}\times\mathbf{a}$.) As scalar functions can be twice differentiated, so can vector functions, for example $\dfrac{\mathrm{d}^2\mathbf{F}}{\mathrm{d}t^2}$.

A vector function can be expressed in component form and then differentiated, and this derivative is the differentiated vector function expressed in component form. This is a consequence of linearity. So

$$\mathbf{F}(t) = F_1(t)\mathbf{i} + F_2(t)\mathbf{j} + F_3(t)\mathbf{k}$$

and

$$\frac{\mathrm{d}\mathbf{F}}{\mathrm{d}t} = \frac{\mathrm{d}F_1}{\mathrm{d}t}\mathbf{i} + \frac{\mathrm{d}F_2}{\mathrm{d}t}\mathbf{j} + \frac{\mathrm{d}F_3}{\mathrm{d}t}\mathbf{k}$$

are the components of its derivative. The alternative notations $\mathbf{F}(t) = (F_1(t), F_2(t), F_3(t))$ and $\dfrac{\mathrm{d}\mathbf{F}}{\mathrm{d}t} = \left(\dfrac{\mathrm{d}F_1}{\mathrm{d}t}, \dfrac{\mathrm{d}F_2}{\mathrm{d}t}, \dfrac{\mathrm{d}F_3}{\mathrm{d}t}\right)$ are also often convenient.

A vector can also be a function of many variables, for example $\mathbf{F}(x, y, z)$. Partial derivatives are defined in precisely the same way as for scalar functions of many variables, for example

$$\frac{\partial\mathbf{F}}{\partial x} = \lim_{\Delta x\to 0}\left\{\frac{\mathbf{F}(x+\Delta x, y, z) - \mathbf{F}(x, y, z)}{\Delta x}\right\}$$

and so on.

There are also special kinds of vector derivative called grad, div and curl which have a chapter of their own (Chapter 7). In this chapter, we will concentrate on applications to differential geometry and mechanics.

Differential Geometry A curve in three dimensions is defined by the vector function

$$\mathbf{r}(t) = \mathbf{i}x(t) + \mathbf{j}y(t) + \mathbf{k}z(t)$$

The vector function $\mathbf{T}(t) = \dfrac{d\mathbf{r}}{dt}$ is in the direction of the *tangent* to the curve. The vector function $\hat{\mathbf{T}} = \dfrac{d\mathbf{r}}{dt} \bigg/ \left|\dfrac{d\mathbf{r}}{dt}\right|$ is therefore the *unit* tangent to the curve ($\left|\dfrac{d\mathbf{r}}{dt}\right| \neq 0$ of course). The vector function $\dfrac{d\hat{\mathbf{T}}}{dt}$ is in a direction called the *principal normal* to the curve. The unit vector in this direction is written $\hat{\mathbf{N}}$. The unit vectors $\hat{\mathbf{T}}$ and $\hat{\mathbf{N}}$, together with $\hat{\mathbf{T}} \times \hat{\mathbf{N}}$ which is written $\hat{\mathbf{B}}$ and called the *binormal*, form a right-handed co-ordinate system at each point of the curve. The following relations are valid. They define, in addition, two scalar quantities κ, the *curvature* and τ, the *torsion*, and they are called the *Serret–Frenet formulae*.

$$\frac{d\hat{\mathbf{T}}}{ds} = \kappa\hat{\mathbf{N}}, \quad \frac{d\hat{\mathbf{N}}}{ds} = \tau\hat{\mathbf{B}} - \kappa\hat{\mathbf{N}}, \quad \frac{d\hat{\mathbf{B}}}{ds} = -\tau\hat{\mathbf{N}}$$

The letter s denotes arc length along the curve and $\left|\dfrac{d\mathbf{r}}{dt}\right| = \dfrac{ds}{dt}$.

Mechanics The vector function $\mathbf{r}(t)$ can represent the position vector of a particle. The vector function $\mathbf{v} = \dfrac{d\mathbf{r}}{dt}$ represents the velocity, and the derivative of $\mathbf{v}$, $\dfrac{d\mathbf{v}}{dt}$, represents the acceleration. The letter t denotes time, and is the natural parameter to describe the position vector of a particle. In terms of the differential geometry of the *path* of the particle, $\mathbf{r} = \mathbf{r}(s)$ and the particle moves along a curve the form of which is dictated by the forces on it. Differentiating $\mathbf{r}$ with respect to t gives

$$\mathbf{v} = \dot{s}\hat{\mathbf{T}}$$
$$\mathbf{a} = \ddot{s}\hat{\mathbf{T}} + \dot{s}^2\kappa\hat{\mathbf{N}}$$

The force on the particle, by Newton's Second Law, is $\mathbf{F}(t) = m\mathbf{a}$.

6.2 Worked Examples

Example 6.1 Prove that, if $\mathbf{F}(t)$ and $\mathbf{G}(t)$ are differentiable vector functions at $t = t_0$, then $\mathbf{F} + \mathbf{G}$ and $\mathbf{F} \times \mathbf{G}$ are also differentiable vector functions.

Solution This example is here to show that all is well when it comes to being able to differentiate sums and products of differentiable vector functions. Since both $\mathbf{F}(t)$ and $\mathbf{G}(t)$ are differentiable at $t = t_0$, the limits

$\mathbf{F}'(t) = \lim\limits_{t\to t_0}\left\{\dfrac{\mathbf{F}(t) - \mathbf{F}(t_0)}{t - t_0}\right\}$ and $\mathbf{G}'(t) = \lim\limits_{t\to t_0}\left\{\dfrac{\mathbf{G}(t) - \mathbf{G}(t_0)}{t - t_0}\right\}$ both exist. The derivative of the sum $\mathbf{F}(t) + \mathbf{G}(t)$ is defined as

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{F}(t)+\mathbf{G}(t)) = \lim_{t\to t_0}\left\{\frac{\mathbf{F}(t)+\mathbf{G}(t)-\mathbf{F}(t_0)-\mathbf{G}(t_0)}{t-t_0}\right\}$$
$$= \lim_{t\to t_0}\left\{\frac{\mathbf{F}(t)-\mathbf{F}(t_0)}{t-t_0}+\frac{\mathbf{G}(t)-\mathbf{G}(t_0)}{t-t_0}\right\}$$
$$= \lim_{t\to t_0}\left\{\frac{\mathbf{F}(t)-\mathbf{F}(t_0)}{t-t_0}\right\}+\lim_{t\to t_0}\left\{\frac{\mathbf{G}(t)-\mathbf{G}(t_0)}{t-t_0}\right\}$$

The right-hand side is $\mathbf{F}'(t)+\mathbf{G}'(t)$, hence $\mathbf{F}(t)+\mathbf{G}(t)$ is differentiable at $t = t_0$. Consider now the cross product $\mathbf{F}(t)\times\mathbf{G}(t)$ in particular the limit

$$\lim_{t\to t_0}\left\{\frac{\mathbf{F}(t)\times\mathbf{G}(t)-\mathbf{F}(t_0)\times\mathbf{G}(t_0)}{t-t_0}\right\}$$
$$= \lim_{t\to t_0}\left\{\frac{\mathbf{F}(t)\times\mathbf{G}(t)-\mathbf{F}(t_0)\times\mathbf{G}(t)+\mathbf{F}(t_0)\times\mathbf{G}(t)-\mathbf{F}(t_0)\times\mathbf{G}(t_0)}{t-t_0}\right\}$$
$$= \lim_{t\to t_0}\left\{\frac{\mathbf{F}(t)-\mathbf{F}(t_0)}{t-t_0}\times\mathbf{G}(t)\right\}+\lim_{t\to t_0}\left\{\frac{\mathbf{G}(t)-\mathbf{G}(t_0)}{t-t_0}\times\mathbf{F}(t_0)\right\}$$
$$= \mathbf{F}'(t)\times\mathbf{G}(t)+\mathbf{F}(t)\times\mathbf{G}'(t)$$

Of course, $\mathbf{F}(t_0)$ and $\mathbf{G}(t_0)$ both exist as, by the information in the question, do their derivatives $\mathbf{F}'(t_0)$ and $\mathbf{G}'(t_0)$. Thus $\mathbf{F}(t)\times\mathbf{G}(t)$ is differentiable with derivative $\mathbf{F}'(t)\times\mathbf{G}(t)+\mathbf{F}(t)\times\mathbf{G}'(t)$ evaluated at $t = t_0$.

Example 6.2 Show that a differentiable vector function $\mathbf{F}(t)$ that is a unit vector has a derivative that is everywhere perpendicular to $\mathbf{F}(t)$ itself.

Solution If $\mathbf{F}(t)$ is a unit vector, then $|\mathbf{F}(t)|^2 = \mathbf{F}(t)\cdot\mathbf{F}(t) = 1$. Differentiating this using the product rule gives $\mathbf{F}'(t)\cdot\mathbf{F}(t)+\mathbf{F}(t)\cdot\mathbf{F}'(t) = 0$ so that $\mathbf{F}'(t)\cdot\mathbf{F}(t) = 0$. This last result indeed implies that $\mathbf{F}(t)$ is everywhere perpendicular to $\mathbf{F}'(t)$. This result is very useful for problems involving geometry (see Example 6.6).

Example 6.3 Show that the vector function $\mathbf{F}(t) = \mathbf{a}\cos\omega t+\mathbf{b}\sin\omega t$ satisfies the differential equation $\mathbf{F}''(t)+\omega^2\mathbf{F}(t) = \mathbf{0}$, and interpret this geometrically. (Assume $\mathbf{a}\times\mathbf{b}\neq\mathbf{0}$.)

Solution Differentiating $\mathbf{F}(t) = \mathbf{a}\cos\omega t+\mathbf{b}\sin\omega t$ gives

$$\mathbf{F}'(t) = \omega\mathbf{a}\sin\omega t+\omega\mathbf{b}\cos\omega t$$

Differentiating again gives

$$\mathbf{F}''(t) = -\omega^2\mathbf{a}\cos\omega t-\omega^2\mathbf{b}\sin\omega t = -\omega^2\mathbf{F}(t)$$

Thus $\mathbf{F}(t)$ satisfies the equation

$$\mathbf{F}''(t)+\omega^2\mathbf{F}(t) = \mathbf{0}$$

Now, writing $\mathbf{a}$ and $\mathbf{b}$ in component form, assuming $\mathbf{i}$ and $\mathbf{j}$ to be in the plane of $\mathbf{a}$ and $\mathbf{b}$ gives $\mathbf{F}(t) = \mathbf{i}(a_1\cos\omega t+b_1\sin\omega t)+\mathbf{j}(a_2\cos\omega t+b_2\sin\omega t)$. In order to make progress we call upon matrix theory. If we write

$$X = a_1\cos\omega t+b_1\sin\omega t$$
$$Y = a_2\cos\omega t+b_2\sin\omega t$$

then, using matrix notation

$$\begin{pmatrix}X\\Y\end{pmatrix} = \begin{pmatrix}a_1 & b_1\\a_2 & b_2\end{pmatrix}\begin{pmatrix}\cos\omega t\\\sin\omega t\end{pmatrix} \tag{1}$$

or, $\mathbf{F} = \mathbf{A}\begin{pmatrix}\cos\omega t\\ \sin\omega t\end{pmatrix}$ where $\mathbf{A}$ stands for the 2×2 matrix $\begin{pmatrix}a_1 & b_1\\ a_2 & b_2\end{pmatrix}$. Elimination of t results in a quadratic expression connecting X and Y which in fact represents an ellipse. This becomes clearer after the following algebra. Since $\mathbf{a} \times \mathbf{b} \neq \mathbf{0}$, vectors $\mathbf{a}$ and $\mathbf{b}$ are not parallel and so the 2×2 matrix $\begin{pmatrix}a_1 & b_1\\ a_2 & b_2\end{pmatrix}$ is non-singular and can be diagonalised. This process involves finding the eigenvalues of the matrix and their associated eigenvectors (see the Fact Sheet, section 4.1, page 50). It is equivalent to a rotation of the axes so that the new set coincides with the principal axes of the ellipse. The matrix equation (1) transforms to

$$\mathbf{x} = \mathbf{PX} = \mathbf{PAP}^{-1}\mathbf{P}\begin{pmatrix}\cos\omega t\\ \sin\omega t\end{pmatrix}$$

where $\mathbf{PAP}^{-1}$ is diagonal. Thus, in terms of new co-ordinates x and y

$$x = k_1 \cos\omega t$$

$$y = k_2 \sin\omega t$$

so that $$\frac{x^2}{k_1^2} + \frac{y^2}{k_2^2} = 1$$ is the ellipse.

As t varies, the co-ordinates (x, y) (and of course (X, Y)) trace out the ellipse, the period of traverse of the ellipse being $\frac{2\pi}{\omega}$. This, as might be guessed, can be related to planetary motion, see Chapter 8 in the author's *Work Out Mechanics*.

Example 6.4 Given the scalar function $\phi = xyz$ and the vector function $\mathbf{A} = x^2\mathbf{i} + y^2\mathbf{j} + z^2\mathbf{k}$ find the following derivatives:

(a) $\frac{\partial}{\partial x}(\phi\mathbf{A})$, (b) $\frac{\partial^2\mathbf{A}}{\partial x^2} + \frac{\partial^2\mathbf{A}}{\partial y^2} + \frac{\partial^2\mathbf{A}}{\partial z^2}$, (c) $\frac{\partial^2}{\partial y\partial z}\left(\frac{\mathbf{A}}{\phi}\right)$,

(d) $\mathbf{A}\cdot\left(\frac{\partial\mathbf{A}}{\partial x} + \frac{\partial\mathbf{A}}{\partial y} + \frac{\partial\mathbf{A}}{\partial z}\right)$.

Solution This example provides some reasonably routine practice in differentiating and manipulating vectors.

(a) From the formulae for ϕ and $\mathbf{A}$, $\phi\mathbf{A} = x^3yz\mathbf{i} + xy^3z\mathbf{j} + xyz^3\mathbf{k}$, so

$\frac{\partial}{\partial x}(\phi\mathbf{A}) = 3x^2yz\mathbf{i} + y^3z\mathbf{j} + yz^3\mathbf{k}$.

(b) Since $\mathbf{A} = x^2\mathbf{i} + y^2\mathbf{j} + z^2\mathbf{k}$, $\frac{\partial^2\mathbf{A}}{\partial x^2} = 2\mathbf{i}$ etc., so

$\frac{\partial^2\mathbf{A}}{\partial x^2} + \frac{\partial^2\mathbf{A}}{\partial y^2} + \frac{\partial^2\mathbf{A}}{\partial z^2} = 2\mathbf{i} + 2\mathbf{j} + 2\mathbf{k}$. We shall see in Chapter 7 that

$\frac{\partial^2\mathbf{A}}{\partial x^2} + \frac{\partial^2\mathbf{A}}{\partial y^2} + \frac{\partial^2\mathbf{A}}{\partial z^2}$ can be written $\nabla^2\mathbf{A}$ and is called the Laplacian.

(c) $\frac{\mathbf{A}}{\phi} = \frac{x}{yz}\mathbf{i} + \frac{y}{xz}\mathbf{j} + \frac{z}{xy}\mathbf{k}$, so $\frac{\partial}{\partial y}\left(\frac{\mathbf{A}}{\phi}\right) = -\frac{x}{y^2z}\mathbf{i} + \frac{1}{xz}\mathbf{j} - \frac{z}{xy^2}\mathbf{k}$ giving

$\frac{\partial^2}{\partial y\partial z}\left(\frac{\mathbf{A}}{\phi}\right) = \frac{x}{y^2z^2}\mathbf{i} - \frac{1}{xz^2}\mathbf{j} - \frac{1}{xy^2}\mathbf{k}$.

(d) Now, $\frac{\partial\mathbf{A}}{\partial x} = 2x\mathbf{i}$, $\frac{\partial\mathbf{A}}{\partial y} = 2y\mathbf{j}$ and $\frac{\partial\mathbf{A}}{\partial z} = 2z\mathbf{k}$, therefore

$$\mathbf{A}\cdot\left(\frac{\partial\mathbf{A}}{\partial x} + \frac{\partial\mathbf{A}}{\partial y} + \frac{\partial\mathbf{A}}{\partial z}\right) = (x^2\mathbf{i} + y^2\mathbf{j} + z^2\mathbf{k})\cdot(2x\mathbf{i} + 2y\mathbf{j} + 2z\mathbf{k}) = 2x^3 + 2y^3 + 2z^3.$$

Example 6.5 If a vector function **A** depends on x, y, z and t, but, in turn x, y and z are functions of t prove that $\dfrac{d\mathbf{A}}{dt} = \dfrac{\partial \mathbf{A}}{\partial t} + \dfrac{\partial \mathbf{A}}{\partial x}\dfrac{dx}{dt} + \dfrac{\partial \mathbf{A}}{\partial y}\dfrac{dy}{dt} + \dfrac{\partial \mathbf{A}}{\partial z}\dfrac{dz}{dt}$.

Solution This example is a version of the chain rule (see Chapter 2) for vectors. However, although its proof is reasonably straightforward – it is achieved by using the components of **A** – it has wide applicability in continuum mechanics, especially fluid mechanics. The components of the vector function $\mathbf{A}(x, y, z, t)$ are defined as usual as

$$\mathbf{A}(x, y, z, t) = A_1(x, y, z, t)\mathbf{i} + A_2(x, y, z, t)\mathbf{j} + A_3(x, y, z, t)\mathbf{k}$$

where of course the components of **A**, $A_1(x, y, z, t)$, $A_2(x, y, z, t)$, $A_3(x, y, z, t)$ are standard scalar functions of four variables. Since x, y and z are functions of t, we have

$$\frac{dA_1}{dt} = \frac{\partial A_1}{\partial t} + \frac{\partial A_1}{\partial x}\frac{dx}{dt} + \frac{\partial A_1}{\partial y}\frac{dy}{dt} + \frac{\partial A_1}{\partial z}\frac{dz}{dt}$$

and similarly for A_2 and A_3. Forming the sum $\mathbf{i}\dfrac{dA_1}{dt} + \mathbf{j}\dfrac{dA_2}{dt} + \mathbf{k}\dfrac{dA_3}{dt}$ which is, of course $\dfrac{d\mathbf{A}}{dt}$ we get $\dfrac{d\mathbf{A}}{dt} = \dfrac{\partial \mathbf{A}}{\partial t} + \dfrac{\partial \mathbf{A}}{\partial x}\dfrac{dx}{dt} + \dfrac{\partial \mathbf{A}}{\partial y}\dfrac{dy}{dt} + \dfrac{\partial \mathbf{A}}{\partial z}\dfrac{dz}{dt}$ as required.

Example 6.6 If s is the arc length along a curve, $\hat{\mathbf{T}}$ is the unit tangent, $\hat{\mathbf{N}}$ the principal normal and $\hat{\mathbf{B}} = \hat{\mathbf{T}} \times \hat{\mathbf{N}}$ is the binormal, show that $\dfrac{d\hat{\mathbf{T}}}{ds} = \kappa\hat{\mathbf{N}}$, $\dfrac{d\hat{\mathbf{B}}}{ds} = -\tau\hat{\mathbf{N}}$ and $\dfrac{d\hat{\mathbf{N}}}{ds} = \tau\hat{\mathbf{B}} - \kappa\hat{\mathbf{T}}$ where κ and τ are constants.

Solution Since $\hat{\mathbf{T}}\cdot\hat{\mathbf{T}} = 1$, we have already shown that $\hat{\mathbf{T}}\cdot\dfrac{d\hat{\mathbf{T}}}{ds} = 0$, that is $\hat{\mathbf{T}}$ is perpendicular to $\dfrac{d\hat{\mathbf{T}}}{ds}$ (see Example 6.2). We now *define* $\hat{\mathbf{N}}$ to be in the direction of $\dfrac{d\hat{\mathbf{T}}}{ds}$, so that $\dfrac{d\hat{\mathbf{T}}}{ds} = \kappa\hat{\mathbf{N}}$. $\hat{\mathbf{N}}$ is in the direction perpendicular to the tangent $\hat{\mathbf{T}}$, but $\hat{\mathbf{N}}$ and $\hat{\mathbf{T}}$ lie in the *plane of the curve* (at the point defined by s). This plane is the only plane of the curve for an infinitesimal length of arc δs, unless we have a plane curve of course, in which case $\hat{\mathbf{T}}$, $\hat{\mathbf{N}}$ and the curve itself all lie in one plane. The constant κ is called the *curvature* of the curve and is, in general, a function of s. If κ is a constant, and the curve is planar, then the curve itself must be a circle. The quantity $\dfrac{1}{\kappa} = \rho$ is called the radius of curvature of the curve and is equal to the radius for a circle. Now consider $\hat{\mathbf{B}} = \hat{\mathbf{T}} \times \hat{\mathbf{N}}$, and differentiate both sides with respect to s to give

$$\frac{d\hat{\mathbf{B}}}{ds} = \hat{\mathbf{T}} \times \frac{d\hat{\mathbf{N}}}{ds} + \frac{d\hat{\mathbf{T}}}{ds} \times \hat{\mathbf{N}}$$

The last term is zero, since $\dfrac{d\hat{\mathbf{T}}}{ds}$ is parallel to $\hat{\mathbf{N}}$. Hence $\dfrac{d\hat{\mathbf{B}}}{ds} = \hat{\mathbf{T}} \times \dfrac{d\hat{\mathbf{N}}}{ds}$, so both $\hat{\mathbf{T}}$ and $\dfrac{d\hat{\mathbf{N}}}{ds}$ are perpendicular to $\dfrac{d\hat{\mathbf{B}}}{ds}$. However, $\hat{\mathbf{T}}$ and $\dfrac{d\hat{\mathbf{N}}}{ds}$ are also perpendicular to $\hat{\mathbf{N}}$ for the following reasons: $\hat{\mathbf{T}}$ by construction, and $\dfrac{d\hat{\mathbf{N}}}{ds}$ because we can differentiate $\hat{\mathbf{N}}\cdot\hat{\mathbf{N}} = 1$ (similarly to $\hat{\mathbf{T}}\cdot\hat{\mathbf{T}} = 1$) and deduce that $\hat{\mathbf{N}}\cdot\dfrac{d\hat{\mathbf{N}}}{ds} = 0$. Thus we conclude that $\hat{\mathbf{N}}$ is parallel to $\dfrac{d\hat{\mathbf{B}}}{ds}$ (they are both perpendicular to the plane of $\hat{\mathbf{T}}$ and $\dfrac{d\hat{\mathbf{N}}}{ds}$). We define τ to be the constant such that

$$\frac{d\hat{\mathbf{B}}}{ds} = -\tau\hat{\mathbf{N}}$$

τ is called the *torsion* of the curve.

$\hat{\mathbf{B}} = \hat{\mathbf{T}} \times \hat{\mathbf{N}}$ indicates that $\hat{\mathbf{T}}$, $\hat{\mathbf{N}}$ and $\hat{\mathbf{B}}$ form a right-handed system of unit vectors, so $\hat{\mathbf{N}} = \hat{\mathbf{B}} \times \hat{\mathbf{T}}$. Differentiating this with respect to s gives

$$\frac{d\hat{\mathbf{N}}}{ds} = \hat{\mathbf{B}} \times \frac{d\hat{\mathbf{T}}}{ds} + \frac{d\hat{\mathbf{B}}}{ds} \times \hat{\mathbf{T}}$$
$$= \hat{\mathbf{B}} \times \kappa\hat{\mathbf{N}} - \tau\hat{\mathbf{N}} \times \hat{\mathbf{T}}$$
$$= -\kappa\hat{\mathbf{T}} + \tau\hat{\mathbf{B}}$$

We have thus deduced the following three equations: $\frac{d\hat{\mathbf{T}}}{ds} = \kappa\hat{\mathbf{N}}$, $\frac{d\hat{\mathbf{B}}}{ds} = -\tau\hat{\mathbf{N}}$ and $\frac{d\hat{\mathbf{N}}}{ds} = \tau\hat{\mathbf{B}} - \kappa\hat{\mathbf{T}}$ where κ and τ are constants. These are called the Serret–Frenet formulae, and they can be put in the more memorable matrix form:

$$\frac{d}{ds}\begin{pmatrix} \hat{\mathbf{T}} \\ \hat{\mathbf{N}} \\ \hat{\mathbf{B}} \end{pmatrix} = \begin{pmatrix} 0 & \kappa & 0 \\ -\kappa & 0 & \tau \\ 0 & -\tau & 0 \end{pmatrix}\begin{pmatrix} \hat{\mathbf{T}} \\ \hat{\mathbf{N}} \\ \hat{\mathbf{B}} \end{pmatrix}$$

In the following two examples, we show how to find $\hat{\mathbf{T}}$, $\hat{\mathbf{N}}$ and $\hat{\mathbf{B}}$ as well as κ and τ for some specific curves, first however here are two observations.

(1) If $\kappa = 0$ then $\hat{\mathbf{T}}$ is a constant which means that the curve must be a straight line. In this case, $\hat{\mathbf{N}}$ and $\hat{\mathbf{B}}$ define the plane of the normal to the line, but the locations of $\hat{\mathbf{N}}$ and $\hat{\mathbf{B}}$ themselves are arbitrary except that $\hat{\mathbf{N}}$ is perpendicular to $\hat{\mathbf{B}}$. The Serret–Frenet formulae are all identically zero in this case.

(2) If $\tau = 0$ then $\hat{\mathbf{B}}$ is a constant, so $\hat{\mathbf{T}}$ and $\hat{\mathbf{N}}$ lie in the same plane for all s. The curve must also lie in this plane, so the curve is planar. The next problem gives an example of such a curve.

Example 6.7 Determine $\hat{\mathbf{T}}$, $\hat{\mathbf{N}}$, $\hat{\mathbf{B}}$, κ and τ for the parabola $y^2 = 4ax$, where a is a constant.

Solution First of all, we parameterise the parabola in terms of a parameter t as follows: $x = at^2$, $y = 2at$. This is by no means unique, but it has the merit of avoiding square roots. Parameterisations can always be checked by eliminating the parameter which, in this instance, regains $y^2 = 4ax$. The next step is to find $\frac{ds}{dt} = \left|\frac{d\mathbf{r}}{dt}\right|$. To do this, note that

$$\mathbf{r} = x\mathbf{i} + y\mathbf{j} + (0\mathbf{k})$$
$$= at^2\mathbf{i} + 2at\mathbf{j}$$

so
$$\frac{d\mathbf{r}}{dt} = 2at\mathbf{i} + 2a\mathbf{j}$$

and
$$\frac{ds}{dt} = (4a^2t^2 + 4a^2)^{1/2} = 2a\sqrt{t^2 + 1}$$

The unit tangent is thus $\hat{\mathbf{T}} = \dfrac{\frac{d\mathbf{r}}{dt}}{\left|\frac{d\mathbf{r}}{dt}\right|} = \dfrac{\frac{d\mathbf{r}}{dt}}{\frac{ds}{dt}}\left(= \dfrac{d\mathbf{r}}{ds}\right) = \dfrac{2at\mathbf{i} + 2a\mathbf{j}}{2a\sqrt{t^2 + 1}}$

$$= \mathbf{i}\frac{t}{\sqrt{t^2 + 1}} + \mathbf{j}\frac{1}{\sqrt{t^2 + 1}} = \frac{t\mathbf{i} + \mathbf{j}}{\sqrt{t^2 + 1}}$$

To find $\hat{\mathbf{N}}$, we differentiate to obtain

$$\frac{d\hat{\mathbf{T}}}{dt} = \mathbf{i}\frac{\sqrt{1 + t^2} - t\frac{1}{2}(1 + t^2)^{-\frac{1}{2}}\,2t}{(1 + t^2)} + \mathbf{j}\frac{-t}{(1 + t^2)^{\frac{3}{2}}}$$
$$= \mathbf{i}\frac{1}{(1 + t^2)^{3/2}} - \mathbf{j}\frac{t}{(1 + t^2)^{3/2}}$$

So
$$\frac{d\hat{\mathbf{T}}}{ds} = \frac{d\hat{\mathbf{T}}}{dt}\bigg/\frac{ds}{dt} = \frac{\mathbf{i} - t\mathbf{j}}{2a(1 + t^2)^2}$$

This quantity is $\kappa\hat{\mathbf{N}}$ by the Serret–Frenet formulae. The modulus of this derivative is thus κ and is given by $\left|\frac{d\hat{\mathbf{T}}}{ds}\right| = \kappa = \dfrac{1}{2a(1 + t^2)^2}[1^2 + t^2]^{\frac{1}{2}} = \dfrac{1}{2a(1 + t^2)^{3/2}}$, so $\hat{\mathbf{N}} = \dfrac{d\hat{\mathbf{T}}}{ds}\bigg/\left|\dfrac{d\hat{\mathbf{T}}}{ds}\right| = \dfrac{\mathbf{i} - t\mathbf{j}}{\sqrt{1 + t^2}}$.

It is often worth checking that $|\hat{\mathbf{T}}| = 1$, $|\hat{\mathbf{N}}| = 1$ and $\hat{\mathbf{T}}\cdot\hat{\mathbf{N}} = 0$ at this stage, especially if the amount of algebraic manipulation is large. The next step is to find $\hat{\mathbf{B}} = \hat{\mathbf{T}} \times \hat{\mathbf{N}}$. Using the determinant for the cross product gives $\hat{\mathbf{B}} = \dfrac{1}{1+t^2}\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ t & 1 & 0 \\ 1 & -t & 0 \end{vmatrix} = \dfrac{1}{1+t^2}\mathbf{k}(-t^2 - 1) = -\mathbf{k}$ which is constant (consistent with the remarks made at the end of the last example; a parabola is a plane curve). Obviously $\tau = 0$ since $\dfrac{d\hat{\mathbf{B}}}{ds} = 0$. Summarising our results, thus

$$\hat{\mathbf{T}} = \frac{\mathbf{i}t + \mathbf{j}}{\sqrt{1+t^2}}, \quad \hat{\mathbf{N}} = \frac{\mathbf{i} - \mathbf{j}t}{\sqrt{1+t^2}}, \quad \hat{\mathbf{B}} = -\mathbf{k}, \; \kappa = \frac{1}{2a(1+t^2)^{\frac{3}{2}}}, \; \tau = 0$$

The next example is three dimensional.

Example 6.8 A curve is parameterised by the following equations: $x(\theta) = \theta\sqrt{2}$, $y(\theta) = e^{\theta}$, $z(\theta) = e^{-\theta}$. Determine $\hat{\mathbf{T}}$, $\hat{\mathbf{N}}$, $\hat{\mathbf{B}}$, κ and τ for this curve.

Solution The method follows precisely that of the last example, except that this time the algebra is more involved since the curve is not planar. Computer algebra is definitely useful for this example. A more convenient notation to adopt is $\mathbf{r} = (x, y, z)$ instead of the more usual $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$. Thus

$$\mathbf{r} = (\theta\sqrt{2}, e^{\theta}, e^{-\theta})$$

whence

$$\frac{d\mathbf{r}}{d\theta} = (\sqrt{2}, e^{\theta}, -e^{-\theta})$$

so

$$\left|\frac{d\mathbf{r}}{d\theta}\right| = (2 + e^{2\theta} + e^{-2\theta})^{\frac{1}{2}}$$

$$= (e^{\theta} + e^{-\theta})$$

$$= 2\cosh\theta$$

Therefore

$$\hat{\mathbf{T}} = \frac{d\mathbf{r}}{ds} = \frac{d\mathbf{r}}{d\theta}\bigg/\frac{ds}{d\theta} = \frac{1}{2\cosh\theta}(\sqrt{2}, e^{\theta}, -e^{-\theta}).$$

We need to differentiate this with respect to θ, and to do this we must use the product formula. The result is a little messy:

$$\frac{d\hat{\mathbf{T}}}{d\theta} = \frac{1}{2\cosh\theta}(0, e^{\theta}, e^{-\theta}) - \frac{\sinh\theta}{2\cosh^2\theta}(\sqrt{2}, e^{\theta}, -e^{-\theta})$$

but in fact this simplifies to

$$\frac{d\hat{\mathbf{T}}}{d\theta} = \frac{1}{2\cosh^2\theta}\left(-\sqrt{2}\sinh\theta, e^{\theta}(\cosh\theta - \sinh\theta), e^{-\theta}(\cosh\theta + \sinh\theta)\right)$$

$$= \frac{1}{2\cosh^2\theta}(-\sqrt{2}\sinh\theta, 1, 1)$$

since

$$\cosh\theta - \sinh\theta = \tfrac{1}{2}(e^{\theta} + e^{-\theta}) - \tfrac{1}{2}(e^{\theta} - e^{-\theta}) = e^{-\theta}$$

and

$$\cosh\theta + \sinh\theta = \tfrac{1}{2}(e^{\theta} + e^{-\theta}) + \tfrac{1}{2}(e^{\theta} - e^{-\theta}) = e^{\theta}$$

so

$$\kappa\hat{\mathbf{N}} = \frac{d\hat{\mathbf{T}}}{ds} = \frac{d\hat{\mathbf{T}}}{d\theta}\bigg/\frac{ds}{d\theta} = \frac{1}{4\cosh^3\theta}(-\sqrt{2}\sinh\theta, 1, 1)$$

giving $\kappa = \dfrac{1}{4\cosh^3\theta}(2\sinh^2\theta + 1 + 1)^{\frac{1}{2}} = \dfrac{1}{4\cosh^3\theta}(2\cosh^2\theta)^{\frac{1}{2}} = \dfrac{\sqrt{2}}{4\cosh^2\theta}$, and

$\hat{\mathbf{N}} = \dfrac{1}{\sqrt{2}\cosh\theta}(-\sqrt{2}\sinh\theta, 1, 1)$. It is worth stopping to check that $\hat{\mathbf{T}}\cdot\hat{\mathbf{N}} = 0$

at this point (the other two checks $|\hat{\mathbf{T}}| = 1$ and $|\hat{\mathbf{N}}| = 1$ are less useful since we have implicitly used these to derive our expressions for $\hat{\mathbf{T}}$ and $\hat{\mathbf{N}}$). We see that $\hat{\mathbf{T}}\cdot\hat{\mathbf{N}} = \dfrac{1}{2\sqrt{2}\cosh^2\theta}(-2\sinh\theta + e^{\theta} - e^{-\theta}) = 0$ since $2\sinh\theta = e^{\theta} - e^{-\theta}$. There is every chance therefore that our $\hat{\mathbf{T}}$ and $\hat{\mathbf{N}}$ are correct. We thus press on and calculate $\hat{\mathbf{B}}$ as follows:

$$\hat{\mathbf{B}} = \hat{\mathbf{T}} \times \hat{\mathbf{N}} = \frac{1}{2\sqrt{2}\cosh^2\theta}\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \sqrt{2} & e^{\theta} & -e^{-\theta} \\ -\sqrt{2}\sinh\theta & 1 & 1 \end{vmatrix}$$

$$= \frac{1}{2\sqrt{2}\cosh^2\theta}(e^{\theta} + e^{-\theta}, \sqrt{2}e^{-\theta}\sinh\theta - \sqrt{2}, \sqrt{2} + \sqrt{2}e^{\theta}\sinh\theta)$$

Using for example, $e^{-\theta}\sinh\theta - 1 = \frac{1}{2}(1 - e^{-2\theta}) - 1 = -\frac{1}{2}(e^{-2\theta} + 1) = -e^{-\theta}\cosh\theta$ in the middle term, this simplifies to

$$\hat{\mathbf{B}} = \frac{1}{2\cosh\theta}(\sqrt{2}, -e^{-\theta}, e^{\theta})$$

Two more checks tell us we are probably right, namely

$$\hat{\mathbf{T}}\cdot\hat{\mathbf{B}} = \frac{1}{4\cosh^2\theta}(2 - 1 - 1) = 0$$

and
$$\hat{\mathbf{N}}\cdot\hat{\mathbf{B}} = \frac{1}{2\sqrt{2}\cosh^2\theta}(-2\sinh\theta + e^{-\theta} - e^{\theta}) = 0$$

It only remains to find τ, and to do this, we differentiate $\hat{\mathbf{B}}$ to obtain

$$\begin{aligned}\frac{d\hat{\mathbf{B}}}{d\theta} &= -\frac{\sinh\theta}{2\cosh^2\theta}(\sqrt{2}, -e^{-\theta}, e^{\theta}) + \frac{1}{2\cosh^2\theta}(0, e^{-\theta}, e^{\theta}) \\ &= \frac{1}{2\cosh^2\theta}\left(-\sqrt{2}\sinh\theta, e^{-\theta}(\cosh\theta + \sinh\theta), e^{\theta}(\cosh\theta - \sinh\theta)\right) \\ &= \frac{1}{2\cosh^2\theta}(-\sqrt{2}\sinh\theta, 1, 1)\end{aligned}$$

using similar simplifications as before. Hence, $\dfrac{d\hat{\mathbf{B}}}{d\theta} = \dfrac{\hat{\mathbf{N}}}{\sqrt{2}\cosh\theta}$ giving

$$\frac{d\hat{\mathbf{B}}}{ds} = \frac{d\hat{\mathbf{B}}}{d\theta}\bigg/\frac{ds}{d\theta} = \frac{\hat{\mathbf{N}}}{2\sqrt{2}\cosh^2\theta}$$

and therefore
$$\tau = -\frac{1}{2\sqrt{2}\cosh^2\theta} = -\frac{\sqrt{2}}{4\cosh^2\theta}$$

Summarising, for the given curve, $\mathbf{r} = (\theta\sqrt{2}, e^{\theta}, e^{-\theta})$ we have

$$\hat{\mathbf{T}} = \frac{1}{2\cosh\theta}(\sqrt{2}, e^{\theta}, -e^{-\theta})$$

$$\hat{\mathbf{N}} = \frac{1}{\sqrt{2}\cosh\theta}(-\sqrt{2}\sinh\theta, 1, 1)$$

$$\hat{\mathbf{B}} = \frac{1}{2\cosh\theta}(\sqrt{2}, -e^{-\theta}, e^{\theta})$$

$$\kappa = \frac{\sqrt{2}}{4\cosh^2\theta}, \quad \tau = -\kappa = -\frac{\sqrt{2}}{4\cosh^2\theta}$$

These five quantities give all the information one should ever need about the curve $\mathbf{r} = (\theta\sqrt{2}, e^{\theta}, e^{-\theta})$.

Example 6.9 Determine equations for the velocity $\mathbf{v}$ and acceleration $\mathbf{a}$ of a particle at position vector $\mathbf{r}$ in terms of the tangent, principal normal, binormal, curvature and torsion of its path.

Solution A little particle mechanics is useful as background here. The position vector $\mathbf{r} = \mathbf{r}(t)$, where t is *time*, gives the position of a particle at time t when under the action of a force. The parabolic flight of a particle projected in air neglecting resistance is a good example. The velocity $\mathbf{v}$ is the rate of change of position with time, therefore $\mathbf{v} = \dfrac{d\mathbf{r}}{dt}$, and acceleration $\mathbf{a}$ is the rate of change of velocity with time hence $\mathbf{a} = \dfrac{d\mathbf{v}}{dt}$. Also using the function of a function rule in one variable, $\mathbf{v} = \dfrac{d\mathbf{r}}{dt} = \dfrac{d\mathbf{r}}{ds}\dfrac{ds}{dt} = \dot{s}\dfrac{d\mathbf{r}}{ds}$ where we have used the notation of a dot over a symbol to denote the time rate of change which is reasonably standard. However, $\hat{\mathbf{T}} = \dfrac{d\mathbf{r}}{ds}$ therefore $\mathbf{v} = \hat{\mathbf{T}}\dot{s}$. Differentiating this again with respect to time gives

$$\mathbf{a} = \hat{\mathbf{T}}\ddot{s} + \frac{d\hat{\mathbf{T}}}{dt}\dot{s}$$

using the product rule (where of course $\ddot{s} = \dfrac{d^2s}{dt^2}$). Now, $\dfrac{d\hat{\mathbf{T}}}{dt} = \dfrac{d\hat{\mathbf{T}}}{ds}\dfrac{ds}{dt} = \kappa\hat{\mathbf{N}}\dfrac{ds}{dt} = \kappa\dot{s}\hat{\mathbf{N}}$ using the Serret–Frenet formulae. Hence, substituting back into the formula for $\mathbf{a}$ we get

$$\mathbf{a} = \ddot{s}\hat{\mathbf{T}} + \kappa\dot{s}^2\hat{\mathbf{N}}$$

This formula reduces to $\mathbf{a} = \ddot{s}\hat{\mathbf{T}}$ for motion in a straight line. Also for motion in a circle, $\dot{s} = U$ = a constant (so that $\ddot{s}$ is zero) and $\mathbf{a} = \kappa U^2\hat{\mathbf{N}}$. This may be familiar to some readers, and may become familiar to more if it is recalled that $\kappa = \dfrac{1}{a}$ where a is the radius of the circle of motion, and $\hat{\mathbf{N}}$ is directed towards the centre so that the acceleration is of magnitude U^2/a towards the centre of the circle. This is called centripetal acceleration (due to 'centrifugal force', see Example 6.11). The beauty of the formula $\mathbf{a} = \ddot{s}\hat{\mathbf{T}} + \kappa\dot{s}^2\hat{\mathbf{N}}$ is its general applicability.

Example 6.10 A smooth narrow tube has the shape of the curve described parametrically as follows: $x = ae^{\theta}\cos\theta$, $y = ae^{\theta}\sin\theta$, $z = \sqrt{2}a(e^{\theta} - 1)$ where a is a constant, and the z-axis is vertically downward. A particle of mass m inside the tube is released from rest at the point for which $\theta = 0$, and slides down the tube under gravity ($mg\mathbf{k}$). Find the magnitude of the reaction of the mass m on the tube.

Solution This question is a combination of mechanics and differential geometry; let us concentrate on the latter first. (The former is emphasised in a similar problem in the author's *Work Out Mechanics*, Example 3.8, page 36, published by Macmillan in 1995.) This is a three-dimensional problem, so we need to find $\hat{\mathbf{T}}$, $\hat{\mathbf{N}}$ and $\hat{\mathbf{B}}$. First of all

$$\frac{d\mathbf{r}}{d\theta} = (ae^{\theta}(\cos\theta - \sin\theta), ae^{\theta}(\cos\theta + \sin\theta), \sqrt{2}ae^{\theta})$$

hence

$$\frac{ds}{d\theta} = \left|\frac{d\mathbf{r}}{d\theta}\right| = (a^2e^{2\theta}(\cos\theta - \sin\theta)^2 + a^2e^{2\theta}(\cos\theta + \sin\theta)^2 + 2a^2e^{2\theta})^{\frac{1}{2}}$$
$$= 2ae^{\theta}$$

and
$$\hat{\mathbf{T}} = \frac{d\mathbf{r}}{ds} = \left(\tfrac{1}{2}(\cos\theta - \sin\theta), \tfrac{1}{2}(\cos\theta + \sin\theta), \frac{\sqrt{2}}{2}\right)$$

Differentiating again with respect to θ gives

$$\frac{d\hat{\mathbf{T}}}{d\theta} = (-\tfrac{1}{2}(\cos\theta + \sin\theta), \tfrac{1}{2}(\cos\theta - \sin\theta), 0)$$

so that
$$\frac{d\hat{\mathbf{T}}}{ds} = \frac{e^{-\theta}}{4a}(-\sin\theta - \cos\theta, \cos\theta - \sin\theta, 0)$$

Taking the modulus gives

$$\kappa = \left|\frac{d\hat{\mathbf{T}}}{ds}\right| = \frac{\sqrt{2}e^{-\theta}}{4a}$$

so that
$$\hat{\mathbf{N}} = \frac{1}{\sqrt{2}}(-\sin\theta - \cos\theta, \cos\theta - \sin\theta, 0)$$

Finally, we need to find $\hat{\mathbf{B}} = \hat{\mathbf{T}} \times \hat{\mathbf{N}}$ so that

$$\hat{\mathbf{B}} = \frac{1}{2\sqrt{2}}\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \cos\theta - \sin\theta & \cos\theta + \sin\theta & \sqrt{2} \\ -\cos\theta - \sin\theta & \cos\theta - \sin\theta & 0 \end{vmatrix}$$
$$= \tfrac{1}{2}(\sin\theta - \cos\theta, -\cos\theta - \sin\theta, \sqrt{2})$$

The torsion τ is not required in this problem. We can now tackle the mechanics part of the problem. Newton's Second Law is *force = mass × acceleration.* The right-hand side has already been derived in Example 6.9 in the form $m(\ddot{s}\hat{\mathbf{T}} + \kappa\dot{s}^2\hat{\mathbf{N}})$, while the left-hand side is $\mathbf{R} + mg\mathbf{k}$ where $\mathbf{R}$ is the reaction of the particle on the tube and $\mathbf{k}$ points downward. In mechanics, the equation of motion is resolved into convenient components before being solved. In this example, the convenient co-ordinate system to use is the natural (intrinsic) co-ordinates of the tube, that is $\hat{\mathbf{T}}$, $\hat{\mathbf{N}}$ and $\hat{\mathbf{B}}$. The reaction $\mathbf{R}$ is perpendicular to $\hat{\mathbf{T}}$, so we can write

$$\mathbf{R} = R_N\hat{\mathbf{N}} + R_B\hat{\mathbf{B}}$$

also note that $\hat{\mathbf{T}} + \hat{\mathbf{B}} = (0, 0, \sqrt{2}) = \mathbf{k}\sqrt{2}$ so

$$\mathbf{k} = \frac{1}{\sqrt{2}}(\hat{\mathbf{T}} + \hat{\mathbf{B}})$$

With these relationships, Newton's Second Law becomes the more useful:

$$(R_N\hat{\mathbf{N}} + R_B\hat{\mathbf{B}}) + \frac{mg}{\sqrt{2}}(\hat{\mathbf{T}} + \hat{\mathbf{B}}) = m(\ddot{s}\hat{\mathbf{T}} + \kappa\dot{s}^2\hat{\mathbf{N}})$$

which in component form gives the three scalar equations

$$\frac{mg}{\sqrt{2}} = m\ddot{s}$$
$$R_N = \kappa\ddot{s}$$
$$R_B + \frac{mg}{\sqrt{2}} = 0$$

The first of these is $\dfrac{d^2s}{dt^2} = \dfrac{g}{\sqrt{2}}$ which integrates to $\dfrac{ds}{dt} = \dfrac{gt}{\sqrt{2}}$ since $\dfrac{ds}{dt} = 0$ when $t = 0$ (the particle starts from rest). Hence we can calculate a relationship between θ and t as follows

$$\frac{d\theta}{dt} = \frac{ds}{dt}\Big/\frac{ds}{d\theta} = \frac{gt}{\sqrt{2}}\Big/ 2ae^{\theta}$$

so that
$$2ae^{\theta}\, d\theta = \frac{gt}{\sqrt{2}}\, dt$$

which, upon integration, gives $2\sqrt{2}ae^{\theta} = \dfrac{gt^2}{2}$ + constant. From which

$$2\sqrt{2}ae^{\theta} = 2\sqrt{2}a + \frac{gt^2}{2} \tag{1}$$

since $\theta = 0$ when $t = 0$. We will need this equation later. From the equation $R_N = m\kappa\dot{s}^2$ we get $R_N = m\,\dfrac{\sqrt{2}e^{-\theta}}{4a}\cdot\dfrac{g^2t^2}{2}$ inserting expressions already derived for κ and $\dot{s}$. Substituting for t^2 in terms of e^{θ} from equation (1) leads to

$$R_N = mg(1 - e^{-\theta})$$

which together with $\quad R_B = -\dfrac{mg}{\sqrt{2}}$ gives

$$|\mathbf{R}|^2 = R_N^2 + R_B^2 = m^2g^2((1 - e^{-\theta})^2 + \tfrac{1}{2})$$

so

$$|\mathbf{R}| = mg[\tfrac{1}{2} + (1 - e^{-\theta})^2]^{\frac{1}{2}}$$

is the required magnitude of the reaction. Some stamina is required in order to see this problem through to its conclusion. Figure 6.1 gives a picture of the tube.

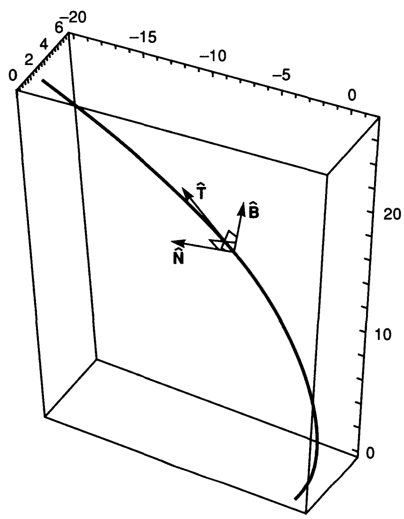


Figure 6.1 The tube – a spiral of expontentially increasing profile with z.

Example 6.11

A two-dimensional co-ordinate system $(\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2)$ rotates at a constant rate $\boldsymbol{\omega}$ about a line which is perpendicular to it. If $\mathbf{i}$ and $\mathbf{j}$ are fixed in the plane, and $\boldsymbol{\omega} = \omega\mathbf{k}$, and if $\hat{\mathbf{e}}_1\cdot\hat{\mathbf{e}}_2 = 0$ show that $\dot{\hat{\mathbf{e}}}_1 = \boldsymbol{\omega}\times\hat{\mathbf{e}}_1$ and $\dot{\hat{\mathbf{e}}}_2 = \boldsymbol{\omega}\times\hat{\mathbf{e}}_2$, hence deduce that the rate of change of any vector quantity $\mathbf{F}$ referred to fixed axes is in fact $\dfrac{d\mathbf{F}}{dt} + \boldsymbol{\omega}\times\mathbf{F}$ referred to $\hat{\mathbf{e}}_1$ and $\hat{\mathbf{e}}_2$. Finally, deduce the form of velocity and acceleration referred to the rotating axes.

Solution

Defining the angle θ as that between $\mathbf{i}$ and $\hat{\mathbf{e}}_1$ at some time t (see Figure 6.2), then $\theta = \omega t$.

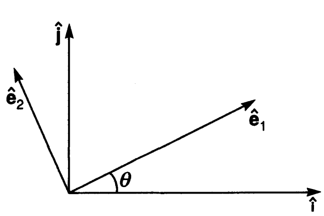


Figure 6.2 The angle $\theta\ (= \omega t)$ between $\hat{\mathbf{i}}$ and $\hat{\mathbf{e}}_1$.

In component form:

$$\hat{\mathbf{e}}_1 = \cos\theta\mathbf{i} + \sin\theta\mathbf{j}$$

and

$$\hat{\mathbf{e}}_2 = -\sin\theta\mathbf{i} + \cos\theta\mathbf{j}$$

With $\theta = \omega t$, we differentiate with respect to t to obtain

$$\frac{d\hat{\mathbf{e}}_1}{dt} = -\omega\sin\omega t\mathbf{i} + \omega\cos\omega t\mathbf{j}$$

and $$\frac{d\hat{\mathbf{e}}_2}{dt} = -\omega\cos\omega t\mathbf{i} + \omega\sin\omega t\mathbf{j}$$

Noting that $\mathbf{k} \times \mathbf{j} = -\mathbf{i}$ and $\mathbf{k} \times \mathbf{i} = \mathbf{j}$ these can be written

$$\begin{aligned}\frac{d\hat{\mathbf{e}}_1}{dt} &= -\omega\sin\omega t(\mathbf{k} \times \mathbf{j}) + \omega\cos\omega t(\mathbf{k} \times \mathbf{i})\\ &= \omega\mathbf{k} \times (\mathbf{j}\sin\omega t + \mathbf{i}\cos\omega t)\\ &= \omega \times \hat{\mathbf{e}}_1\end{aligned}$$

and similarly

$$\begin{aligned}\frac{d\hat{\mathbf{e}}_2}{dt} &= -\omega\cos\omega t(\mathbf{k} \times \mathbf{j}) - \omega\sin\omega t(\mathbf{k} \times \mathbf{i})\\ &= \omega\mathbf{k} \times (\mathbf{j}\cos\omega t - \mathbf{i}\sin\omega t)\\ &= \omega \times \hat{\mathbf{e}}_2\end{aligned}$$

So $\dot{\hat{\mathbf{e}}}_1 = \omega \times \hat{\mathbf{e}}_1$ and $\dot{\hat{\mathbf{e}}}_2 = \omega \times \hat{\mathbf{e}}_2$ as required.

Any vector $\mathbf{F}$ can be referred to the rotating vectors $\hat{\mathbf{e}}_1$ and $\hat{\mathbf{e}}_2$ together with the vector $\mathbf{k}$, that is, $\mathbf{F} = F_1\hat{\mathbf{e}}_1 + F_2\hat{\mathbf{e}}_2 + F_3\mathbf{k}$. Hence

$$\begin{aligned}\frac{d\mathbf{F}}{dt} &= \frac{d}{dt}(F_1\hat{\mathbf{e}}_1 + F_2\hat{\mathbf{e}}_2 + F_3\mathbf{k})\\ &= \frac{dF_1}{dt}\hat{\mathbf{e}}_1 + F_1\frac{d\hat{\mathbf{e}}_1}{dt} + \frac{dF_2}{dt}\hat{\mathbf{e}}_2 + F_2\frac{d\hat{\mathbf{e}}_2}{dt} + \frac{dF_3}{dt}\mathbf{k}\\ &= \left.\frac{d\mathbf{F}}{dt}\right|_f + F_1\frac{d\hat{\mathbf{e}}_1}{dt} + F_2\frac{d\hat{\mathbf{e}}_2}{dt}\\ &= \left.\frac{d\mathbf{F}}{dt}\right|_f + F_1\boldsymbol{\omega} \times \hat{\mathbf{e}}_1 + F_2\boldsymbol{\omega} \times \hat{\mathbf{e}}_2 = \left.\frac{d\mathbf{F}}{dt}\right|_f + \boldsymbol{\omega} \times \mathbf{F}\end{aligned}$$

where the right-hand side is referred to rotating axes. This is indicated by the symbol $\left.\frac{d\mathbf{F}}{dt}\right|_f$. Also note that the term $\boldsymbol{\omega} \times \omega\mathbf{k}$ is zero in the expansion of $\boldsymbol{\omega} \times \mathbf{F}$ in component form. Letting $\mathbf{F}$ be the position vector $\mathbf{r}$ gives immediately

$$\mathbf{v} = \mathbf{v}_f + \boldsymbol{\omega} \times \mathbf{r}$$

where $\mathbf{v}$ is the true velocity, and $\mathbf{v}_f$ is the velocity measured as if $\hat{\mathbf{e}}_1$ and $\hat{\mathbf{e}}_2$ were fixed. Similarly, letting $\mathbf{F}$ be the true velocity vector $\mathbf{v}$ we have on double application of the formula $\frac{d\mathbf{F}}{dt} = \left.\frac{d\mathbf{F}}{dt}\right|_f + \boldsymbol{\omega} \times \mathbf{F}$

$$\mathbf{a} = \left(\left.\frac{d}{dt}\right|_f + \boldsymbol{\omega} \times\right)\left(\left.\frac{d\mathbf{r}}{dt}\right|_f + \boldsymbol{\omega} \times \mathbf{r}\right)$$

Now, $\left.\frac{d\mathbf{r}}{dt}\right|_f = \mathbf{v}_f$, hence

$$\mathbf{a} = \frac{d}{dt}\left(\left.\frac{d\mathbf{r}}{dt}\right|_f + \boldsymbol{\omega} \times \mathbf{r}\right) + \boldsymbol{\omega} \times \left(\left.\frac{d\mathbf{r}}{dt}\right|_f + \boldsymbol{\omega} \times \mathbf{r}\right)$$

Multiplying out, omitting the subscript f as understood, and rearranging gives

$$\begin{aligned}\mathbf{a} &= \frac{d^2\mathbf{r}}{dt^2} + \frac{d\boldsymbol{\omega}}{dt} \times \mathbf{r} + 2\boldsymbol{\omega} \times \frac{d\mathbf{r}}{dt} + \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r})\\ &= \mathbf{a}_f + \dot{\boldsymbol{\omega}} \times \mathbf{r} + 2\boldsymbol{\omega} \times \mathbf{v}_f + \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r})\end{aligned}$$

The $\mathbf{a}_f$ in this equation is the acceleration measured as if $\hat{\mathbf{e}}_1$ and $\hat{\mathbf{e}}_2$ are fixed. The term $\dot{\boldsymbol{\omega}} \times \mathbf{r}$ (often zero) is due to uneven rotation. The term $2\boldsymbol{\omega} \times \mathbf{v}_f$ is called the Coriolis acceleration and plays a central role in geophysical problems. Finally, the term $\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r})$ is the centripetal acceleration and is the generalisation of the '$-\omega^2 r$' term often wrongly called 'centrifugal force' in school mechanics texts.

6.3 Exercises

6.1. If $\mathbf{F}$ is a differentiable vector function of t, and ϕ is a differentiable scalar function of t, show from first principles that

$$\frac{\mathrm{d}}{\mathrm{d}t}(\phi\mathbf{F}) = \phi\frac{\mathrm{d}\mathbf{F}}{\mathrm{d}t} + \mathbf{F}\frac{\mathrm{d}\phi}{\mathrm{d}t}$$

6.2. If F is the magnitude of the differentiable vector function $\mathbf{F}$, both functions of t, then show that

$$\mathbf{F}\cdot\frac{\mathrm{d}\mathbf{F}}{\mathrm{d}t} = F\frac{\mathrm{d}F}{\mathrm{d}t}$$

6.3. In mechanics, $\mathbf{F}$ is a force, m is mass (assumed constant here), $\boldsymbol{\Gamma} = \mathbf{r}\times\mathbf{F}$ is torque and $\mathbf{L} = \mathbf{r}\times m\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}$ is angular momentum. Newton's Second Law states that $\mathbf{F} = m\frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}t^2}$. Show that $\boldsymbol{\Gamma} = \frac{\mathrm{d}\mathbf{L}}{\mathrm{d}t}$, that is torque is equal to the rate of change of angular momentum. Hence deduce that if $\mathbf{F}$ is parallel to $\mathbf{r}$ then angular momentum is a constant.

6.4. A particle moves along the curve described by the position vector

$$\mathbf{r} = (t^2 - 6t)\mathbf{i} + (t^3 - t^2)\mathbf{j}\sqrt{2} + 2t\mathbf{k}.$$

Find the velocity $\mathbf{v}$ and acceleration $\mathbf{a}$ at general time t, and find a specific value of t for which $\mathbf{v}$ and $\mathbf{a}$ are perpendicular.

6.5. Find the derivatives of the expressions

$$\frac{\mathbf{r}+\mathbf{a}}{\mathbf{r}^2+\mathbf{a}^2};\quad \frac{\mathbf{r}\times\mathbf{a}}{\mathbf{r}\cdot\mathbf{a}}.$$

6.6. Find all the values of $\mathbf{r}$ that satisfy the equations
(a) $\frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}t^2} = \mathbf{a}$; (b) $\mathbf{a}\times\frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}t^2} = \mathbf{b}$ $(\mathbf{a}\cdot\mathbf{b} = 0)$.

6.7. Find an expression for the unit tangents of the following curves:
(a) $\mathbf{r} = (t^2, t^3 - t, 0)$,
(b) $\mathbf{r} = (a\sin t, a\cos t, bt)$ (a helix),
(c) $\mathbf{r} = (t, t^2, t^3)$.

6.8. Show that the two curves

$$\mathbf{r} = (1+\lambda, 1+2\lambda, 1+\lambda)$$
$$\mathbf{r} = (2\mu, \mu, 2-4\mu)$$

where λ and μ are arbitrary parameters are straight lines that intersect at right angles.

6.9. The position vector $\mathbf{r}$ satisfies the differential equation $\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = (y, -x, b)$ where x and y are functions of t. Show that this equation is satisfied by the helix of Exercise 6.7(b). (The *vector field* $\mathbf{F} = \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}$ is called the *direction field*. Two-dimensional direction fields are best drawn on a computer screen.)

6.10. Show that for any plane curve, the torsion is zero.

6.11. Consider the helix given by $\mathbf{r} = (a\sin t, a\cos t, bt)$, $0 \le t < 2\pi$. By differentiating and being mindful of definitions (see Example 6.8) find the unit tangent $\hat{\mathbf{T}}$, the principal normal $\hat{\mathbf{N}}$, the binormal $\hat{\mathbf{B}}$, curvature κ and torsion τ for this helix.

6.12. By first showing that $\frac{\mathrm{d}s}{\mathrm{d}\theta} = 3(2\theta^2 + 1)$ where s is the arc length, find the quantities $\hat{\mathbf{T}}$, $\hat{\mathbf{N}}$, $\hat{\mathbf{B}}$, κ and τ for the twisted cubic curve given parametrically by $\mathbf{r} = (2\theta^3, 3\theta^2, 3\theta)$. Show in particular that for this curve, the curvature and torsion have the same magnitude.

6.13. For the curve $\mathbf{r} = (2a\cos t, 2a\sin t, bt^2)$ show that $\dot{\mathbf{r}}\ddot{\mathbf{r}} = 4b^2t$ and $|\dot{\mathbf{r}}\times\ddot{\mathbf{r}}| = 4a(a^2 + b^2 + b^2t^2)^{\frac{1}{2}}$, where the dot denotes differentiation with respect to t, and determine $\hat{\mathbf{T}}$.

6.14. If the position vector of a mass m is given by $\mathbf{r} = \left(\frac{U}{\omega}[1 - \cos\omega t], \frac{U}{\omega}\sin\omega t, Vt\right)$ where U and V are constants then show that $m\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = q(\mathbf{v}\times\mathbf{B})$ where $\mathbf{v} = \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}$ and $\mathbf{B} = (0, 0, B) =$ a constant vector, and q is a constant. [This is the problem of a mass m with charge q moving in a magnetic field $(0, 0, B)$.]

6.15. If $\mathbf{r} = (x, y, z)$ then show that $\kappa = \sqrt{(x'')^2 + (y'')^2 + (z'')^2}$ where the dash denotes differentiation with respect to the arc length s.

6.16. Show that $\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}\cdot\frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}s^2}\times\frac{\mathrm{d}^3\mathbf{r}}{\mathrm{d}s^3} = -\tau\kappa^2$ for any curve described by the position vector $\mathbf{r}$.

6.17. Use the chain rule to show that if $\mathbf{r} = \mathbf{r}(u, v)$ represents a surface (u and v are parameters) then $\mathrm{d}s^2 = E\mathrm{d}u^2 + 2F\mathrm{d}u\mathrm{d}v + G\mathrm{d}v^2$ where $E = \frac{\partial\mathbf{r}}{\partial u}\cdot\frac{\partial\mathbf{r}}{\partial u}$, $F = \frac{\partial\mathbf{r}}{\partial u}\cdot\frac{\partial\mathbf{r}}{\partial v}$ and $G = \frac{\partial\mathbf{r}}{\partial v}\cdot\frac{\partial\mathbf{r}}{\partial v}$. Hence prove that a necessary and sufficient condition that the curvilinear co-ordinates u, v be orthogonal is $F = 0$.

Topic Guide

Definitions
Properties
Vector Identities
Laplacian
Cylindrical and Spherical Co-ordinates

7 Gradient, Divergence, Curl and Curvilinear Co-ordinates

7.1 Fact Sheet

There are three 'vector differential operators', grad, div and curl. The definition of these in Cartesian co-ordinates is:

$$\text{grad } \phi = \nabla\phi = \hat{\mathbf{i}}\frac{\partial\phi}{\partial x} + \hat{\mathbf{j}}\frac{\partial\phi}{\partial y} + \hat{\mathbf{k}}\frac{\partial\phi}{\partial z},$$ where ϕ is a scalar once differentiable function of x, y, z.

$$\text{div } \mathbf{F} = \nabla\cdot\mathbf{F} = \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z},$$ where $\mathbf{F}$ is a once differentiable vector function of x, y, z.

$$\text{curl } \mathbf{F} = \nabla\times\mathbf{F} = \hat{\mathbf{i}}\left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}\right) + \hat{\mathbf{j}}\left(\frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}\right) + \hat{\mathbf{k}}\left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right)$$

$$= \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_1 & F_2 & F_3 \end{vmatrix}$$

The symbol ∇ (called del) represents the vector operator $\hat{\mathbf{i}}\frac{\partial}{\partial x} + \hat{\mathbf{j}}\frac{\partial}{\partial y} + \hat{\mathbf{k}}\frac{\partial}{\partial z}$. The notation 'grad', 'div' and 'curl' is favoured in UK texts, whereas the notation $\nabla\phi$, $\nabla\cdot\mathbf{F}$, $\nabla\times\mathbf{F}$ is favoured by USA texts. This text uses both but favours the USA notation. Note that $\nabla\phi$ is del operating on a *scalar function* ϕ to produce a *vector*, $\nabla\cdot\mathbf{F}$ is del operating on a *vector function* $\mathbf{F}$ to produce a *scalar*, $\nabla\times\mathbf{F}$ is del operating on a *vector function* $\mathbf{F}$ to produce a *vector*.

The directional derivative is a function $\phi(x, y, z)$ in the direction $\hat{\boldsymbol{\alpha}}$ and is defined as

$$\frac{\partial\phi}{\partial\alpha} = \lim_{\overrightarrow{PP'}\to 0}\left(\frac{\phi_P - \phi_{P'}}{\overrightarrow{PP'}}\right)$$

where ϕ_P and $\phi_{P'}$ are the values of ϕ at points P and P' respectively and $\overrightarrow{PP'}$ is in the direction of $\hat{\boldsymbol{\alpha}}$.

For computational purposes, the relationship

$$\frac{\partial\phi}{\partial\alpha} = \hat{\boldsymbol{\alpha}}\cdot\nabla\phi$$

is useful. The following vector identities are also useful:

$$\nabla\cdot(\phi\mathbf{F}) = \nabla\phi\cdot\mathbf{F} + \phi(\nabla\mathbf{F})$$
$$\nabla\times(\phi\mathbf{F}) = \nabla\phi\times\mathbf{F} + \phi(\nabla\times\mathbf{F})$$
$$\nabla(\mathbf{F}\times\mathbf{G}) = \mathbf{G}\cdot\nabla\times\mathbf{F} - \mathbf{F}\cdot\nabla\times\mathbf{G}$$
$$\nabla\times(\mathbf{F}\times\mathbf{G}) = (\mathbf{G}\cdot\nabla)\mathbf{F} - \mathbf{G}(\nabla\cdot\mathbf{F}) - (\mathbf{F}\cdot\nabla)\mathbf{G} + \mathbf{F}(\nabla\cdot\mathbf{G})$$

where $\mathbf{F} \cdot \nabla = F_1\dfrac{\partial}{\partial x} + F_2\dfrac{\partial}{\partial y} + F_3\dfrac{\partial}{\partial z}$ is a (scalar) differential operator.

$$\nabla(\mathbf{F} \cdot \mathbf{G}) = (\mathbf{G} \cdot \nabla)\mathbf{F} + (\mathbf{F} \cdot \nabla)\mathbf{G} + \mathbf{G} \times (\nabla \times \mathbf{F}) + \mathbf{F} \times (\nabla \times \mathbf{G})$$

$$\nabla \cdot (\nabla\phi) = \nabla^2\phi = \frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2}$$

The operator ∇^2 is called the Laplacian.

$$\begin{aligned} \nabla \times (\nabla\phi) &\equiv \mathbf{0} \quad && \text{curl of gradient} = \mathbf{0} \\ \nabla \cdot (\nabla \times \mathbf{F}) &\equiv 0 && \text{divergence of curl} = 0 \\ \nabla \times (\nabla \times \mathbf{F}) &= \nabla(\nabla \cdot \mathbf{F}) - \nabla^2\mathbf{F} && \end{aligned}$$

If (u, v, w) is an orthogonal curvilinear co-ordinate system, then $x = x(u, v, w)$, $y = y(u, v, w)$ and $z = z(u, v, w)$ so $\mathbf{r} = \mathbf{r}(u, v, w)$.

The unit vectors in the directions u = const., v = const. and w = const. are $\hat{\mathbf{e}}_1$, $\hat{\mathbf{e}}_2$, $\hat{\mathbf{e}}_3$ where

$$\hat{\mathbf{e}}_1 = \frac{\partial \mathbf{r}}{\partial u} \Big/ \left|\frac{\partial \mathbf{r}}{\partial u}\right|, \; \hat{\mathbf{e}}_2 = \frac{\partial \mathbf{r}}{\partial v} \Big/ \left|\frac{\partial \mathbf{r}}{\partial v}\right|, \; \hat{\mathbf{e}}_3 = \frac{\partial \mathbf{r}}{\partial w} \Big/ \left|\frac{\partial \mathbf{r}}{\partial w}\right|$$

and $\left|\dfrac{\partial \mathbf{r}}{\partial u}\right|$, $\left|\dfrac{\partial \mathbf{r}}{\partial v}\right|$ and $\left|\dfrac{\partial \mathbf{r}}{\partial w}\right|$ are written h_1, h_2 and h_3 respectively and characterise the curvilinear system. For cylindrical polar coordinates (R, θ, z) h_1, h_2, h_3 are 1, R, 1 and for spherical polar co-ordinates they are 1, r, $r\sin\theta$ – see Example 7.19.

7.2 Worked Examples

Example 7.1 Show that the vector grad$\phi = \nabla\phi$ defined by

$$\nabla\phi = \hat{\mathbf{i}}\,\frac{\partial\phi}{\partial x} + \hat{\mathbf{j}}\,\frac{\partial\phi}{\partial y} + \hat{\mathbf{k}}\,\frac{\partial\phi}{\partial z}$$

is everywhere perpendicular to the surface $\phi(x, y, z)$ = const.

Solution Let C_1 and C_2 be two curves completely embedded in the surface $\phi(x, y, z)$ = const. Let C_1 be represented by the parametric equations $x = x(t)$, $y = y(t)$, $z = z(t)$ so that on C_1, $\mathbf{r} = (x(t), y(t), z(t))$. Similarly, let C_2 be represented by the parametric equation $x = x(s)$, $y = y(s)$, $z = z(s)$ so that, on C_2, $\mathbf{r} = (x(s), y(s), z(s))$. Now, $\dfrac{d\mathbf{r}}{dt} = \left(\dfrac{dx}{dt}, \dfrac{dy}{dt}, \dfrac{dz}{dt}\right)$ is tangent to C_1; to see this recall that the definition of $\dfrac{d\mathbf{r}}{dt}$ is as follows:

$$\frac{d\mathbf{r}}{dt} = \lim_{\Delta t \to 0}\left\{\frac{\mathbf{r}(t + \Delta t) - \mathbf{r}(t)}{\Delta t}\right\}$$

and with $\Delta\mathbf{r} = \mathbf{r}(t + \Delta t) - \mathbf{r}(t)$ being the third side of the triangle as shown in Figure 7.1 in the limit, $\Delta\mathbf{r}$ must become a tangent to the curve C_1. Therefore, $\dfrac{d\mathbf{r}}{dt}$ must also be tangent to the curve C_1. Similarly, $\dfrac{d\mathbf{r}}{ds} = \left(\dfrac{dx}{ds}, \dfrac{dy}{ds}, \dfrac{dz}{ds}\right)$ is tangent to C_2. Consider the scalar product $\nabla\phi \cdot \dfrac{d\mathbf{r}}{dt}$; this is $\nabla\phi \cdot \dfrac{d\mathbf{r}}{dt} = \dfrac{\partial\phi}{\partial x}\dfrac{dx}{dt} + \dfrac{\partial\phi}{\partial y}\dfrac{dy}{dt} + \dfrac{\partial\phi}{\partial z}\dfrac{dz}{dt}$. However, the right-hand side is $\dfrac{d\phi}{dt}$ by the chain rule (see Chapter 2) and since both $\nabla\phi$ and $\dfrac{d\mathbf{r}}{dt}$ are evaluated on the surface ϕ = constant, $\dfrac{d\phi}{dt} = 0$. Thus $\nabla\phi \cdot \dfrac{d\mathbf{r}}{dt} = 0$

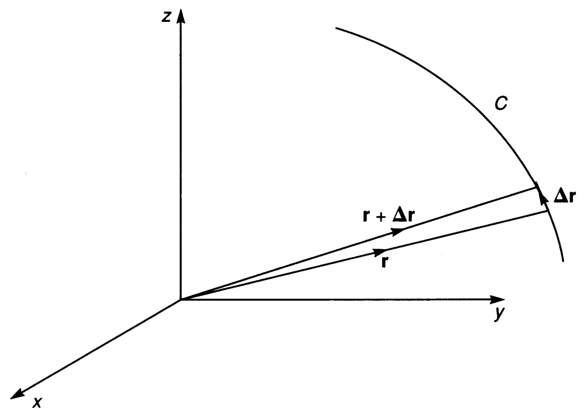


Figure 7.1 The infinitisimal $\mathbf{\Delta r}$ is ultimately tangential to the curve C.

on ϕ = constant. Since neither $\nabla\phi$ nor $\dfrac{d\mathbf{r}}{dt}$ is zero, $\nabla\phi$ is at right angles to $\dfrac{d\mathbf{r}}{dt}$. By a similar argument, $\nabla\phi$ and $\dfrac{d\mathbf{r}}{ds}$ are also at right angles to each other.

Both $\dfrac{d\mathbf{r}}{dt}$ and $\dfrac{d\mathbf{r}}{ds}$, being tangents to the curves embedded in ϕ = constant, are also tangents to the surface ϕ = constant. Thus together $\dfrac{d\mathbf{r}}{dt}$ and $\dfrac{d\mathbf{r}}{ds}$ define the tangent plane at the point of intersection of C_1 and C_2 (see Figure 7.2).

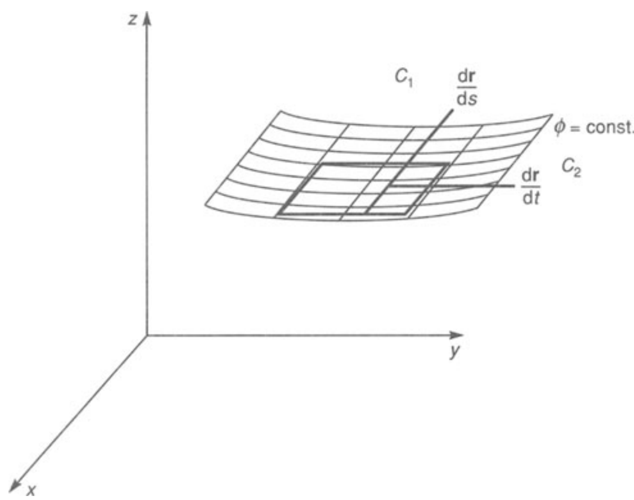


Figure 7.2 The surface ϕ = constant, the embedded curves C_1 and C_2, and the tangents $\dfrac{d\mathbf{r}}{ds}$ and $\dfrac{d\mathbf{r}}{dt}$.

Since $\nabla\phi$ is perpendicular to both $\dfrac{d\mathbf{r}}{dt}$ and $\dfrac{d\mathbf{r}}{ds}$ it must be in the direction of the normal to the surface ϕ = constant.

Example 7.2 Find the unit normal to the surface $x^2y + y^2z + z^2x = k$ = constant at the points (1, 1, 1) and (1, 0, 2). Investigate the behaviour of this surface at the origin (0, 0, 0) in the case where the constant $k = 0$.

Solution Since the normal is in the direction of $\nabla\phi$, we compute the partial derivatives $\dfrac{\partial\phi}{\partial x} = 2xy + z^2$, $\dfrac{\partial\phi}{\partial y} = x^2 + 2yz$, $\dfrac{\partial\phi}{\partial z} = y^2 + 2zx$. Note that the constant k does not feature in any of these derivatives. This is because different values of the constant k merely represent parallel surfaces, the normals of which are therefore also parallel. At (1, 1, 1), $\dfrac{\partial\phi}{\partial x} = 3$, $\dfrac{\partial\phi}{\partial y} = 3$, $\dfrac{\partial\phi}{\partial z} = 3$. Hence $\nabla\phi = 3\mathbf{i} + 3\mathbf{j} + 3\mathbf{k}$ and the unit normal $\dfrac{\nabla\phi}{|\nabla\phi|}$ is found by dividing this by the quantity $\sqrt{3^2 + 3^2 + 3^2} = 3\sqrt{3}$, that is

$\hat{\mathbf{n}} = \dfrac{1}{\sqrt{3}}(\mathbf{i} + \mathbf{j} + \mathbf{k})$. At $(1, 0, 2)$, $\dfrac{\partial\phi}{\partial x} = 9$, $\dfrac{\partial\phi}{\partial y} = 1$, $\dfrac{\partial\phi}{\partial z} = 4$, therefore $\nabla\phi = 9\mathbf{i} + \mathbf{j} + 4\mathbf{k}$ so

$$\hat{\mathbf{n}} = \frac{1}{4\sqrt{6}}(9\mathbf{i} + \mathbf{j} + 4\mathbf{k})$$

When the surface is $x^2y + y^2z + z^2x = 0$, it passes through $(0, 0, 0)$, the origin. Unfortunately, all the derivatives of $\phi(x, y, z) = x^2y + y^2z + z^2x$ are also zero at the origin. Such a surface is said to be *pathological* at $(0, 0, 0)$ and a normal does not exist there. Some notion of the behaviour of the surface $x^2y + y^2z + z^2x = 0$ can be ascertained by considering its intersection with the plane $z = a$. This occurs wherever $x^2y + y^2a + a^2x = 0$. If this is considered as a curve in the x–y plane with a as a parameter to be assigned values, then for large x the curve approximates to the parabola $ay = -x^2$. However, $x > a(4)^{1/3}$ otherwise y is imaginary, yet for small x, $y^2 \approx -ax$. As $a \to 0$, the origin becomes the site of a 'wrinkle', and tangents and normals become ill defined.

Example 7.3 If $\mathbf{F} = \nabla\phi$, and $\mathbf{F} = y^2z^3\mathbf{i} + 2xyz^3\mathbf{j} + 3xy^2z^2\mathbf{k}$ determine ϕ to within an arbitrary constant.

Solution Since $\mathbf{F} = \nabla\phi = \dfrac{\partial\phi}{\partial x}\mathbf{i} + \dfrac{\partial\phi}{\partial y}\mathbf{j} + \dfrac{\partial\phi}{\partial z}\mathbf{k}$, equating this to the given expression for $\mathbf{F}$ leads to $\dfrac{\partial\phi}{\partial x} = y^2z^3$, $\dfrac{\partial\phi}{\partial y} = 2xyz^3$, $\dfrac{\partial\phi}{\partial z} = 3xy^2z^2$. When integrating a *partial* derivative, the arbitrary 'constant' can be a function of those variables that are being held constant under the partial differentiation. Hence when $\dfrac{\partial\phi}{\partial x} = y^2z^3$ is integrated, we obtain $\phi = xy^2z^3 + f_1(y, z)$ where $f_1(y, z)$ is an arbitrary function of y and z. Integrating the expressions for $\dfrac{\partial\phi}{\partial y}$ and $\dfrac{\partial\phi}{\partial z}$ in turn lead to

$$\phi = xy^2z^3 + f_2(y, z)$$
$$\phi = xy^2z^3 + f_3(y, z)$$

Comparing these with each other leads to the conclusion that

$$\phi = xy^2z^3 + c$$

that is, $f_1(y, z) = f_2(x, z) = f_3(x, y) = c =$ constant. We will see later that it is not always possible to find such a function ϕ.

Example 7.4 Establish that if $\phi = x^n + y^n + z^n$ then $\mathbf{r} \cdot \nabla\phi = n\phi$.

Solution If $\phi = x^n + y^n + z^n$, then $\dfrac{\partial\phi}{\partial x} = nx^{n-1}$, $\dfrac{\partial\phi}{\partial y} = ny^{n-1}$, $\dfrac{\partial\phi}{\partial z} = nz^{n-1}$, so $\nabla\phi = n(\mathbf{i}x^{n-1} + \mathbf{j}y^{n-1} + \mathbf{k}z^{n-1})$ and $\mathbf{r} \cdot \nabla\phi = n(x^n + y^n + z^n) = n\phi$ as required.

Example 7.5 Show that $\nabla(\mathbf{a} \cdot \mathbf{r}) = \mathbf{a}$ for any constant vector $\mathbf{a}$.

Solution Suppose that $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$ so that $\mathbf{a} \cdot \mathbf{r} = a_1x + a_2y + a_3z$.

Since $\dfrac{\partial}{\partial x}(\mathbf{a} \cdot \mathbf{r}) = a_1$, $\dfrac{\partial}{\partial y}(\mathbf{a} \cdot \mathbf{r}) = a_2$, $\dfrac{\partial}{\partial z}(\mathbf{a} \cdot \mathbf{r}) = a_3$, we have $\nabla(\mathbf{a} \cdot \mathbf{r}) = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k} = \mathbf{a}$ as required.

Example 7.6 If $\mathbf{E} = -\nabla\phi$, and $\phi = \dfrac{Q}{4\pi kr}$ determine $\mathbf{E}$.

Solution In this problem, $\mathbf{r}$ is the position vector so $r^2 = x^2 + y^2 + z^2$. Using the definition of $\nabla\phi$, we need to calculate the three partial derivatives $\dfrac{\partial\phi}{\partial x}$, $\dfrac{\partial\phi}{\partial y}$ and $\dfrac{\partial\phi}{\partial z}$. So $\dfrac{\partial\phi}{\partial x} = \dfrac{Q}{4\pi k}\dfrac{\partial}{\partial x}(x^2 + y^2 + z^2)^{-\frac{1}{2}}$ $= -\dfrac{Qx}{4\pi k}(x^2 + y^2 + z^2)^{-\frac{3}{2}} = -\dfrac{Qx}{4\pi kr^3}$. By symmetry, $\dfrac{\partial\phi}{\partial y} = -\dfrac{Qy}{4\pi kr^3}$, and $\dfrac{\partial\phi}{\partial z} = -\dfrac{Qz}{4\pi kr^3}$. Thus $\mathbf{E} = \dfrac{Q}{4\pi kr^3}(\mathbf{i}x + \mathbf{j}y + \mathbf{k}z) = \dfrac{Q\mathbf{r}}{4\pi kr^3}$. In fact $\mathbf{E}$ is the electric field due to a single point charge for which $\phi = \dfrac{Q}{4\pi kr}$ is the electric potential.

Example 7.7 Show that, if ϕ is a scalar function of a vector variable $\mathbf{r}$, then

$$\phi(\mathbf{r} + \mathrm{d}\mathbf{r}) = \phi(\mathbf{r}) + (\mathrm{d}\mathbf{r} \cdot \nabla\phi)|_{\mathbf{r}} + O(\mathrm{d}\mathbf{r}^2)$$

where $\mathrm{d}\mathbf{r}$ is a small increment in $\mathbf{r}$. Hence show that the directional derivative in the direction of unit vector $\hat{\boldsymbol{\alpha}}$ is given by $\hat{\boldsymbol{\alpha}} \cdot \nabla\phi$.

Solution Writing $\mathbf{r} = (x, y, z)$ is the convenient notation for this problem, so $\phi(\mathbf{r})$ is in turn written as $\phi(x, y, z)$ whence

$$\phi(\mathbf{r} + \mathrm{d}\mathbf{r}) = \phi(x + \mathrm{d}x, y + \mathrm{d}y, z + \mathrm{d}z)$$

Using Taylor's Series for three variables, we write this as

$$\phi(x + \mathrm{d}x, y + \mathrm{d}y, z + \mathrm{d}z) = \phi(x, y, z) + \left(\mathrm{d}x\frac{\partial\phi}{\partial x} + \mathrm{d}y\frac{\partial\phi}{\partial y} + \mathrm{d}z\frac{\partial\phi}{\partial z}\right)\bigg|_{(x,y,z)} + O(\mathrm{d}x^2, \mathrm{d}y^2, \mathrm{d}z^2)$$

However, $\mathrm{d}\mathbf{r} \cdot \nabla\phi = \mathrm{d}x\dfrac{\partial\phi}{\partial x} + \mathrm{d}y\dfrac{\partial\phi}{\partial y} + \mathrm{d}z\dfrac{\partial\phi}{\partial z}$, whence

$$\phi(\mathbf{r} + \mathrm{d}\mathbf{r}) = \phi(\mathbf{r}) + (\mathrm{d}\mathbf{r} \cdot \nabla\phi)|_{\mathbf{r}} + O(\mathrm{d}\mathbf{r}^2)$$

where all terms on the right-hand side are evaluated at $\mathbf{r}$.

If we align the position vector $\mathbf{r}$ to be in the direction of $\hat{\boldsymbol{\alpha}}$, then

$$\phi(\boldsymbol{\alpha} + \delta\boldsymbol{\alpha}) = \phi(\boldsymbol{\alpha}) + (\delta\boldsymbol{\alpha} \cdot \nabla\phi)_{\boldsymbol{\alpha}} + O(\delta\boldsymbol{\alpha}^2)_{\boldsymbol{\alpha}}$$

so

$$\frac{\phi(\boldsymbol{\alpha} + \delta\boldsymbol{\alpha}) - \phi(\boldsymbol{\alpha})}{\delta\boldsymbol{\alpha}} = \frac{(\mathrm{d}\mathbf{r} \cdot \nabla\phi)_{\boldsymbol{\alpha}}}{\delta\boldsymbol{\alpha}} + O(\delta\boldsymbol{\alpha})_{\boldsymbol{\alpha}}$$

In the limit as $\delta\alpha \to 0$, the left-hand side becomes $\dfrac{\partial\phi}{\partial\alpha}$ which is the directional derivative of the function ϕ in the direction of $\hat{\boldsymbol{\alpha}}$. The right-hand side becomes $\hat{\boldsymbol{\alpha}} \cdot \nabla\phi$ since $\left|\dfrac{\partial\mathbf{r}}{\partial\alpha}\right| = 1$ and $\dfrac{\partial\mathbf{r}}{\partial\alpha}$ is in the direction of $\hat{\boldsymbol{\alpha}}$. Thus to evaluate a directional derivative, we simply find $\nabla\phi$, and take its scalar product with the *unit* vector in the desired direction ($\hat{\boldsymbol{\alpha}}$). Finally insert the values of x, y, z corresponding to the given point. The next Example 7.8 shows the procedure.

Example 7.8 A scalar function ϕ is given by $\phi(x, y, z) = xy^2 - 3xyz + y^3 + z^3$. Determine $\nabla\phi$ at the point (1, 1, 1) and hence determine the directional derivative of ϕ at this point in the direction of $3\mathbf{i} - 2\mathbf{j} + 4\mathbf{k}$.

Solution We use the method hinted at in the last example:

$$\nabla\phi = \mathbf{i}\frac{\partial}{\partial x}(xy^2 - 3xyz + y^3 + z^3) + \mathbf{j}\frac{\partial}{\partial y}(xy^2 - 3xyz + y^3 + z^3) + \mathbf{k}\frac{\partial}{\partial z}(xy^2 - 3xyz + y^3 + z^3)$$

so $\nabla\phi = \mathbf{i}(y^2 - 3yz) + \mathbf{j}(2xy - 3xz + 3y^2) + \mathbf{k}(-3xy + 3z^2)$, and at (1, 1, 1) $\nabla\phi = -2\mathbf{i} + 2\mathbf{j} + 0\mathbf{k}$. In order to find the directional derivative, we need to determine the scalar product of $\nabla\phi$ at

(1,1,1) with the *unit* vector in the direction of 3**i** − 2**j** + 4**k**, whence we calculate

$$(-2, 2, 0) \cdot \frac{1}{\sqrt{29}}(3, -2, 4) = \frac{1}{\sqrt{29}}(-6 - 4 + 0) = -\frac{10}{\sqrt{29}}$$

Example 7.9 A paraboloid of revolution has the equation $5z = x^2 + y^2$. Find the unit normal to the surface at the point (1, 3, 2). Hence obtain the equation of the normal and the tangent plane at that point.

Solution The normal to the surface will be in the direction of the gradient $\nabla(x^2 + y^2 - 5z)$ so $\mathbf{n} = 2x\mathbf{i} + 2y\mathbf{j} - 5\mathbf{k}$. The unit normal at (1, 3, 2) is thus $\dfrac{2\mathbf{i} + 6\mathbf{j} - 10\mathbf{k}}{\sqrt{2^2 + 6^2 + 10^2}} = \dfrac{1}{\sqrt{35}}(\mathbf{i} + 3\mathbf{j} - 5\mathbf{k})$. We now require the equation of the straight line in the direction of $\mathbf{i} + 3\mathbf{j} - 5\mathbf{k}$ through the point (1, 3, 2). In general, the equation of the line parallel to $\mathbf{n} = n_1\mathbf{i} + n_2\mathbf{j} + n_3\mathbf{k}$ through the point (x_1, y_1, z_1) is given by $\dfrac{x - x_1}{n_1} = \dfrac{y - y_1}{n_2} = \dfrac{z - z_1}{n_3}$ (see Example 5.9).

Also the equation of the tangent plane is $(\mathbf{r} - \mathbf{r}_1) \cdot \mathbf{n}$ where $\mathbf{r}_1 = (x_1, y_1, z_1)$ (see Example 5.22). Thus, using the specific point (1, 3, 2) and normal $\mathbf{n} = \mathbf{i} + 3\mathbf{j} - 5\mathbf{k}$ gives the line as $\dfrac{x - 1}{1} = \dfrac{y - 3}{3} = \dfrac{z - 2}{-5}$, and the equation of the plane as

$$(1)(x - 1) + (3)(y - 3) + (-5)(z - 2) = 0$$

which simplifies to $$x + 3y - 5z = 0.$$

Example 7.10 Show that, if **u** is a vector field that represents a fluid velocity, then $\nabla \cdot \mathbf{u}$, the divergence of **u**, represents the amount of fluid created in an arbitrary point in the fluid.

Solution Let Figure 7.3 represent an arbitrary volume in the form of a cuboid of sides δx, δy and δz. One corner of this cuboid is the point (x, y, z) which is our arbitrary point mentioned in the question. Let the velocity **u** have components (u_1, u_2, u_3) so that $\mathbf{u}(x, y, z) = \mathbf{i}u_1(x, y, z) + \mathbf{j}u_2(x, y, z) + \mathbf{k}u_3(x, y, z)$.

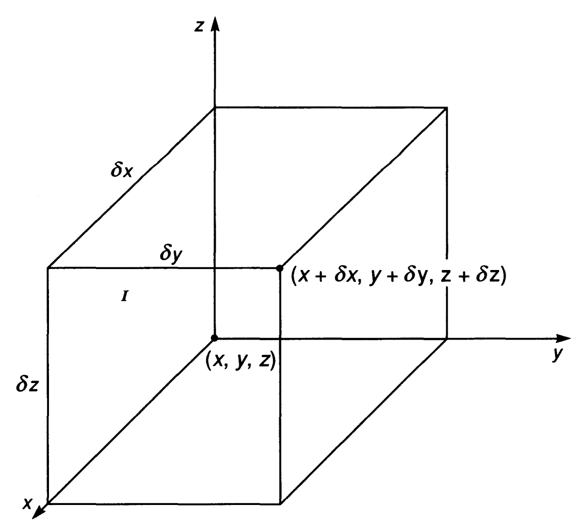


Figure 7.3 An arbitrary cuboid of fluid.

The left-hand side of the cuboid (labelled I in Figure 7.3) has sides δx, δz and so has area $\delta x \delta z$. Since this face is small, the flow through it is approximately

$$u_2(x + \tfrac{1}{2}\delta x, y, z + \tfrac{1}{2}\delta z)$$

as the point $(x + \frac{1}{2}\delta x, y, z + \frac{1}{2}\delta z)$ is precisely at the centre of this face, and u_2 will not differ significantly from its value at the centre throughout the face I. Hence, to this degree of approximation, the flow (read *flux*) of fluid through this face is

$$u_2(x + \tfrac{1}{2}\delta x, y, z + \tfrac{1}{2}\delta z)\delta x\delta y \qquad (A)$$

and this flux is *into* the cuboid. By the same argument, the flux *out of* the cuboid through the opposite face is

$$u_2(x + \tfrac{1}{2}\delta x, y + \delta y, z + \tfrac{1}{2}\delta z)\delta x\delta z \qquad (B)$$

Since, using Taylor's Series (in the one variable y)

$$u_2(x + \tfrac{1}{2}\delta x, y + \delta y, z + \tfrac{1}{2}\delta z) = u_2(x + \tfrac{1}{2}\delta x, y, z + \tfrac{1}{2}\delta z) + \frac{\partial u_2}{\partial y}\delta y + \cdots$$

where $\dfrac{\partial u_2}{\partial y}$ is evaluated at the point $(x + \tfrac{1}{2}\delta x, y, z + \tfrac{1}{2}\delta z)$, the total amount of fluid created by the flow in the y direction is $(B) - (A)$ or $\dfrac{\partial u_2}{\partial y}\delta x\delta y\delta z$ to lowest order in δx, δy and δz.

By considering the other two pairs of opposite faces, we can deduce that the total amount of fluid created inside the cuboid is

$$\left(\frac{\partial u_1}{\partial x} + \frac{\partial u_2}{\partial y} + \frac{\partial u_3}{\partial z}\right)\delta x\delta y\delta z$$

Of course, $\dfrac{\partial u_1}{\partial z}$ is evaluated at $(x, y + \tfrac{1}{2}\delta y, z + \tfrac{1}{2}\delta z)$, $\dfrac{\partial u_2}{\partial y}$ is evaluated at $(x + \tfrac{1}{2}\delta x, y, z + \tfrac{1}{2}\delta z)$ and $\dfrac{\partial u_3}{\partial z}$ is evaluated at $(x + \tfrac{1}{2}\delta x, y + \tfrac{1}{2}\delta y, z)$. However, all of these points tend to (x, y, z) as δx, δy and δz tend to zero. Thus the amount of fluid created *per unit volume* is thus $\left(\dfrac{\partial u_1}{\partial x} + \dfrac{\partial u_2}{\partial y} + \dfrac{\partial u_3}{\partial z}\right)$ which is $\nabla \cdot \mathbf{u}$ and it is evaluated at (x, y, z). This answers the question, giving a physical interpretation for the divergence operator. If $\nabla \cdot \mathbf{u} = 0$ then this expresses the fact that the quantity $\mathbf{u}$ is neither created nor destroyed, that is it is *conserved.* Such an equation is very common in the application of vectors to physical systems.

Example 7.11 Find the value of $\nabla \cdot \mathbf{F}$ for the following vector fields $\mathbf{F}$:
(a) $\mathbf{F} = \mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$
(b) $\mathbf{F} = 2x^2y\mathbf{i} - 2(xy^2 + y^3z)\mathbf{j} + 3y^2z^2\mathbf{k}$
(c) $\mathbf{F} = \mathbf{a}\sin(xyz)$ where $\mathbf{a}$ is a constant vector.

Solution

(a) $\nabla \cdot \mathbf{r} = \dfrac{\partial x}{\partial x} + \dfrac{\partial y}{\partial y} + \dfrac{\partial z}{\partial z} = 1 + 1 + 1 = 3$

(b) $\nabla \cdot \mathbf{F} = \dfrac{\partial}{\partial x}(2x^2y) + \dfrac{\partial}{\partial y}(-2xy^2 - 2y^3z) + \dfrac{\partial}{\partial z}(3y^2z^2)$
$= 4xy - 4xy - 6y^2z + 6y^2z = 0$

Thus $\nabla \cdot \mathbf{F} = 0$. Such a field is called *solenoidal.*

(c) $\nabla \cdot \mathbf{F} = \dfrac{\partial}{\partial x}(a_1\sin xyz) + \dfrac{\partial}{\partial y}(a_2\sin xyz) + \dfrac{\partial}{\partial z}(a_3\sin xyz)$
$= yza_1 \cos xyz + xza_2 \cos xyz + xya_3 \cos xyz$

Note that since $\nabla(\sin xyz) = \mathbf{i}yz \cos xyz + \mathbf{j}xz \cos xyz + \mathbf{k}xy \cos xyz$,
$\mathbf{a} \cdot \nabla(\sin xyz) = a_1yz \cos xyz + a_2xz \cos xyz + a_3xy \cos xyz$
Hence $\nabla \cdot (\mathbf{a}\sin xyz) = \mathbf{a} \cdot \nabla(\sin xyz)$. This is an example of a *vector identity*, we return to these in Example 7.14 (see also Appendix B).

Example 7.12 Given that $\phi = 3xy^2z^3$ find $\nabla \cdot (\nabla\phi)$. Show that $\nabla \cdot (\nabla\phi) = \nabla^2\phi$ and verify this result for the given scalar function ϕ.

Solution If $\phi = 3xy^2z^3$, $\nabla\phi = 3y^2z^3\mathbf{i} + 6xyz^3\mathbf{j} + 9xy^2z^2\mathbf{k}$ hence

$$\nabla \cdot (\nabla\phi) = \frac{\partial}{\partial x}(3y^2z^3) + \frac{\partial}{\partial y}(6xyz^3) + \frac{\partial}{\partial z}(9xy^2z^2)$$

$$= 0 + 6xz^3 + 18xy^2z = 6xz^3 + 18xy^2z$$

In general $\nabla \cdot (\nabla\phi) = \nabla \cdot \left(\mathbf{i}\dfrac{\partial\phi}{\partial x} + \mathbf{j}\dfrac{\partial\phi}{\partial y} + \mathbf{k}\dfrac{\partial\phi}{\partial z}\right)$

$$= \frac{\partial}{\partial x}\left(\frac{\partial\phi}{\partial x}\right) + \frac{\partial}{\partial y}\left(\frac{\partial\phi}{\partial y}\right) + \frac{\partial}{\partial z}\left(\frac{\partial\phi}{\partial z}\right)$$

$$= \frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2} = \nabla^2\phi \text{ (by definition)}$$

With $\phi = 3xy^2z^3$, $\nabla^2\phi = \dfrac{\partial^2}{\partial x^2}(3xy^2z^3) + \dfrac{\partial^2}{\partial y^2}(3xy^2z^3) + \dfrac{\partial^2}{\partial z^2}(3xy^2z^3)$

$$= 0 + 6xz^3 + 18xy^2z = 6xz^3 + 18xy^2z$$

as previously obtained.

Example 7.13 Calculate $\nabla \times \mathbf{F}$ (curl $\mathbf{F}$) where $\mathbf{F}$ is given by the following expressions:
(a) $\mathbf{F} = \boldsymbol{\omega} \times \mathbf{r}$ ($\boldsymbol{\omega}$ = constant) representing the rotation of a rigid body.
(b) $\mathbf{F} = \mathbf{i}yz + \mathbf{j}zx + \mathbf{k}xy$
(c) $\mathbf{F} = \nabla\phi$, where ϕ is any scalar function with continuous second-order partial derivatives.

Solution For all of these expressions for $\mathbf{F} = (F_1, F_2, F_3)$ we use the definition

$$\nabla \times \mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ F_1 & F_2 & F_3 \end{vmatrix} = \mathbf{i}\left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}\right) + \mathbf{j}\left(\frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}\right) + \mathbf{k}\left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right)$$

(a) If $\mathbf{F} = \boldsymbol{\omega} \times \mathbf{r}$, then $\mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \omega_1 & \omega_2 & \omega_3 \\ x & y & z \end{vmatrix}$ and so

$F_1 = \omega_2 z - \omega_3 y$, $\quad F_2 = \omega_3 x - \omega_1 z$, $\quad F_3 = \omega_1 y - \omega_2 x$. Hence

$$\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z} = \omega_1 - (-\omega_1) = 2\omega_1$$

$$\frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x} = \omega_2 - (-\omega_2) = 2\omega_2$$

$$\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} = \omega_3 - (-\omega_3) = 2\omega_3$$

and thus $\nabla \times \mathbf{F} = 2\omega_1\mathbf{i} + 2\omega_2\mathbf{j} + 2\omega_3\mathbf{k} = 2\boldsymbol{\omega}$.
(b) With $\mathbf{F} = \mathbf{i}yz + \mathbf{j}zx + \mathbf{k}xy$

$$\nabla \times \mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ yz & zx & xy \end{vmatrix} = \mathbf{i}(x - x) + \mathbf{j}(y - y) + \mathbf{k}(z - z) = \mathbf{0}$$

Thus $\nabla \times \mathbf{F} = \mathbf{0}$. Such a field is called *irrotational*. More will be said about irrotational fields later.

(c) If $\mathbf{F} = \nabla\phi$ then $F_1 = \dfrac{\partial\phi}{\partial x}$, $F_2 = \dfrac{\partial\phi}{\partial y}$, $F_3 = \dfrac{\partial\phi}{\partial z}$, where $\mathbf{F} = (F_1, F_2, F_3)$ in component form. Thus

$$\nabla \times \mathbf{F} = \mathbf{i}\left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}\right) + \mathbf{j}\left(\frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}\right) + \mathbf{k}\left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right)$$

$$= \mathbf{i}\left(\frac{\partial^2 \phi}{\partial y \partial z} - \frac{\partial^2 \phi}{\partial z \partial y}\right) + \mathbf{j}\left(\frac{\partial^2 \phi}{\partial z \partial x} - \frac{\partial^2 \phi}{\partial x \partial z}\right) + \mathbf{k}\left(\frac{\partial^2 \phi}{\partial x \partial y} - \frac{\partial^2 \phi}{\partial y \partial x}\right)$$

which is zero provided as stated in the question the scalar function ϕ has continuous second-order partial derivatives. Hence whenever $\mathbf{F} = \nabla\phi$, $\nabla \times \mathbf{F} = \mathbf{0}$, that is $\mathbf{F}$ is irrotational. It is also true that if $\nabla \times \mathbf{F} = \mathbf{0}$ then there exists a scalar function ϕ such that $\mathbf{F} = \nabla\phi$. If $\nabla \times \mathbf{F} = \mathbf{0}$ then $\mathbf{F}$ is also called *conservative*, and the scalar ϕ is called the *potential function* (or potential energy function). We shall return to this important topic in Chapter 8 (see Example 8.3). See also Appendix A on conjugate functions.

Example 7.14 Establish the following vector identities:
(a) $\nabla \cdot (\mathbf{F}\phi) = \mathbf{F} \cdot \nabla\phi + \phi\nabla \cdot \mathbf{F}$,
(b) $\nabla \cdot (\mathbf{F} \times \mathbf{G}) = \mathbf{G} \cdot \nabla \times \mathbf{F} - \mathbf{F} \cdot \nabla \times \mathbf{G}$,
(c) $\nabla \times \nabla\phi \equiv \mathbf{0}$,
(d) $\nabla \cdot (\nabla \times \mathbf{F}) \equiv 0$.
Where ϕ is a scalar function that is twice differentiable, and $\mathbf{F}$ and $\mathbf{G}$ are vector functions that are twice differentiable.

Solution The general way to verify identities involving grad, div and curl is to express all the vectors in component form and use the standard rules for the differentiation of a product.

(a) Writing $\mathbf{F} = \mathbf{i}F_1(x, y, z) + \mathbf{j}F_2(x, y, z) + \mathbf{k}F_3(x, y, z)$,

$$\phi\mathbf{F} = \mathbf{i}\phi F_1(x, y, z) + \mathbf{j}\phi F_2(x, y, z) + \mathbf{k}\phi F_3(x, y, z) \text{ and hence}$$

$$\nabla \cdot (\phi\mathbf{F}) = \frac{\partial}{\partial x}(\phi F_1) + \frac{\partial}{\partial y}(\phi F_2) + \frac{\partial}{\partial z}(\phi F_3)$$

$$= \phi\frac{\partial F_1}{\partial x} + F_1\frac{\partial \phi}{\partial x} + \phi\frac{\partial F_2}{\partial y} + F_2\frac{\partial \phi}{\partial y} + \phi\frac{\partial F_3}{\partial z} + F_3\frac{\partial \phi}{\partial z}$$

$$\text{So } \nabla \cdot (\phi\mathbf{F}) = \phi\left(\frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z}\right) + F_1\frac{\partial \phi}{\partial x} + F_2\frac{\partial \phi}{\partial y} + F_3\frac{\partial \phi}{\partial z}$$

$$= \phi\nabla \cdot \mathbf{F} + \mathbf{F} \cdot \nabla\phi \text{ as required.}$$

(b) A similar approach is used for this problem. The components of $\mathbf{F}$ and $\mathbf{G}$ are $\mathbf{F} = \mathbf{i}F_1 + \mathbf{j}F_2 + \mathbf{k}F_3$ and $\mathbf{G} = \mathbf{i}G_1 + \mathbf{j}G_2 + \mathbf{k}G_3$, thus

$$\mathbf{F} \times \mathbf{G} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ F_1 & F_2 & F_3 \\ G_1 & G_2 & G_3 \end{vmatrix} = \mathbf{i}(F_2G_3 - F_3G_2) + \mathbf{j}(F_3G_1 - F_1G_3) + \mathbf{k}(F_1G_2 - F_2G_1)$$

$$\text{so } \nabla \cdot (\mathbf{F} \times \mathbf{G}) = \frac{\partial}{\partial x}(F_2G_3 - F_3G_2) + \frac{\partial}{\partial y}(F_3G_1 - F_1G_3) + \frac{\partial}{\partial z}(F_1G_2 - F_2G_1).$$

$$\text{Now, } \frac{\partial}{\partial x}(F_2G_3 - F_3G_2) = F_2\frac{\partial G_3}{\partial x} + G_3\frac{\partial F_2}{\partial x} - F_3\frac{\partial G_2}{\partial x} - G_2\frac{\partial F_3}{\partial x}$$

and similarly for the other two derivatives. Gathering the twelve terms together in the following convenient order gives:

$$\begin{aligned}\nabla \cdot (\mathbf{F} \times \mathbf{G}) = {} & F_2\frac{\partial G_3}{\partial x} + G_3\frac{\partial F_2}{\partial x} - F_3\frac{\partial G_2}{\partial x} - G_2\frac{\partial F_3}{\partial x} \\ & + F_3\frac{\partial G_1}{\partial y} + G_1\frac{\partial F_3}{\partial y} - F_1\frac{\partial G_3}{\partial y} - G_3\frac{\partial F_1}{\partial y} \\ & + F_1\frac{\partial G_2}{\partial z} + G_2\frac{\partial F_1}{\partial z} - F_2\frac{\partial G_1}{\partial z} - G_1\frac{\partial F_2}{\partial z}\end{aligned}$$

or

$$\nabla \cdot (\mathbf{F} \times \mathbf{G}) = G_1\frac{\partial F_3}{\partial y} - G_1\frac{\partial F_2}{\partial z} + G_2\frac{\partial F_1}{\partial z} - G_2\frac{\partial F_3}{\partial x}$$

$$+ G_3\frac{\partial F_2}{\partial x} - G_3\frac{\partial F_1}{\partial y} - F_1\frac{\partial G_3}{\partial y} + F_1\frac{\partial G_2}{\partial z}$$

$$- F_2\frac{\partial G_1}{\partial z} + F_2\frac{\partial G_3}{\partial x} - F_3\frac{\partial G_2}{\partial x} - F_3\frac{\partial G_1}{\partial y}$$

The right-hand side is $\mathbf{G} \cdot \nabla \times \mathbf{F} - \mathbf{F} \cdot \nabla \times \mathbf{G}$, as required.

(c) Since $\nabla\phi = \mathbf{i}\frac{\partial \phi}{\partial x} + \mathbf{j}\frac{\partial \phi}{\partial y} + \mathbf{k}\frac{\partial \phi}{\partial z}$, we have

$$\nabla \times \nabla\phi = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \frac{\partial \phi}{\partial x} & \frac{\partial \phi}{\partial y} & \frac{\partial \phi}{\partial z} \end{vmatrix} = \mathbf{i}\left(\frac{\partial}{\partial y}\left(\frac{\partial \phi}{\partial z}\right) - \frac{\partial}{\partial z}\left(\frac{\partial \phi}{\partial y}\right)\right) + \text{two similar terms.}$$

All components are thus zero provided that

$$\frac{\partial}{\partial y}\left(\frac{\partial \phi}{\partial z}\right) = \frac{\partial}{\partial z}\left(\frac{\partial \phi}{\partial y}\right), \quad \frac{\partial}{\partial x}\left(\frac{\partial \phi}{\partial y}\right) = \frac{\partial}{\partial y}\left(\frac{\partial \phi}{\partial x}\right) \text{ and } \frac{\partial}{\partial x}\left(\frac{\partial \phi}{\partial z}\right) = \frac{\partial}{\partial z}\left(\frac{\partial \phi}{\partial x}\right)$$

which is equivalent to demanding that ϕ has continuous second-order partial derivatives.

(d) Again we have that

$$\nabla \times \mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_1 & F_2 & F_3 \end{vmatrix} = \mathbf{i}\left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}\right) + \mathbf{j}\left(\frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}\right) + \mathbf{k}\left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right)$$

Hence $\nabla \cdot (\nabla \times \mathbf{F}) = \frac{\partial}{\partial x}\left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}\right) + \frac{\partial}{\partial y}\left(\frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}\right) + \frac{\partial}{\partial z}\left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right)$

$$= \frac{\partial^2 F_3}{\partial x \partial y} - \frac{\partial^2 F_3}{\partial y \partial x} + \frac{\partial^2 F_1}{\partial y \partial z} - \frac{\partial^2 F_1}{\partial z \partial y} + \frac{\partial^2 F_2}{\partial z \partial x} - \frac{\partial^2 F_2}{\partial x \partial z} = 0$$

provided all the components of $\mathbf{F}$ have continuous second-order partial derivatives.

Example 7.15 Determine the value of $\nabla \times (\mathbf{r}f(r))$, where $\mathbf{r}$ is the position vector.

Solution Using the formula for curl, we have

$$\nabla \times (\mathbf{r}f(r)) = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ xf(r) & yf(r) & zf(r) \end{vmatrix}$$

Before expanding this, note that, for example

$$\frac{\partial f}{\partial y} = \frac{\mathrm{d}f}{\mathrm{d}r}\frac{\partial r}{\partial y} = f'\frac{\partial}{\partial y}\sqrt{x^2 + y^2 + z^2} = f' \cdot \frac{1}{2}(x^2 + y^2 + z^2)^{-\frac{1}{2}} \cdot 2y = f'\frac{y}{r}$$

so that $\frac{\partial f}{\partial x} = f'\frac{x}{r}$ and $\frac{\partial f}{\partial z} = f'\frac{z}{r}$. This means that the expression

$$z\frac{\partial f}{\partial y} - y\frac{\partial f}{\partial z} = zf'\frac{y}{r} - yf'\frac{z}{r} = 0$$

This is the first component of $\nabla \times (\mathbf{r}f(r))$. Similarly the other two components are zero. Hence $\nabla \times (\mathbf{r}f(r)) = \mathbf{0}$. In Example 7.20 we meet a quicker method, but this involves knowledge of curvilinear co-ordinates (in particular spherical polars).

Example 7.16 A transformation from Cartesian co-ordinates (x, y, z) to a curvilinear system of co-ordinates (u, v, w). Find expressions for the unit vectors $\hat{\mathbf{e}}_1$, $\hat{\mathbf{e}}_2$ and $\hat{\mathbf{e}}_3$ along the directions given by u varying (v and w constant), v varying (w and u constant) and w varying (u and v constant).

Solution The position vector is $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ where, in terms of u, v and w:

$$x = x(u, v, w),\ y = y(u, v, w) \text{ and } z = z(u, v, w)$$

If u_0, v_0 and w_0 are fixed values of u, v and w respectively, then the vectors $\mathbf{r} = \mathbf{r}(u, v_0, w_0)$, $\mathbf{r} = \mathbf{r}(u_0, v, w_0)$, $\mathbf{r} = \mathbf{r}(u_0, v_0, w)$ represent lines along each of the curvilinear co-ordinates (u, v, w). To see this, note that

$$\begin{aligned} \mathbf{r}(u, v_0, w_0) &= x(u, v_0, w_0)\mathbf{i} + y(u, v_0, w_0)\mathbf{j} + z(u, v_0, w_0)\mathbf{k} \\ \mathbf{r}(u_0, v, w_0) &= x(u_0, v, w_0)\mathbf{i} + y(u_0, v, w_0)\mathbf{j} + z(u_0, v, w_0)\mathbf{k} \\ \text{and } \mathbf{r}(u_0, v_0, w) &= x(u_0, v_0, w)\mathbf{i} + y(u_0, v_0, w)\mathbf{j} + z(u_0, v_0, w)\mathbf{k} \end{aligned}$$

These are all one parameter vectors, and in much the same way as $\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k}$ is the parametric equation of a curve, so are each of the above with u, v and w in turn taking on the role of t. In our case, the three curves are the curves of intersection of the three co-ordinate surfaces u = const., v = const. and w = const., and are of course the three co-ordinate axes.

The derivative of $\mathbf{r}(u, v_0, w_0)$ (with respect to u of course) will be a tangent to this particular co-ordinate curve. This derivative is $\dfrac{\partial \mathbf{r}}{\partial u}$ and the unit tangent is thus $\dfrac{\partial \mathbf{r}}{\partial u}\Big/\left|\dfrac{\partial \mathbf{r}}{\partial u}\right|$. Now the transformation from (x, y, z) to (u, v, w) is guaranteed non-singular provided the Jacobian $\dfrac{\partial(x, y, z)}{\partial(u, v, w)}$ is not zero (see Chapter 2). Writing this in determinant form:

$$\begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} & \dfrac{\partial x}{\partial w} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} & \dfrac{\partial y}{\partial w} \\ \dfrac{\partial z}{\partial u} & \dfrac{\partial z}{\partial v} & \dfrac{\partial z}{\partial w} \end{vmatrix}$$

we see that the components of $\dfrac{\partial \mathbf{r}}{\partial u}$, $\dfrac{\partial \mathbf{r}}{\partial v}$ and $\dfrac{\partial \mathbf{r}}{\partial w}$ are the columns. Hence the modulus of each of the tangents is not zero. Defining $h_1 = \left|\dfrac{\partial \mathbf{r}}{\partial u}\right|$, $h_2 = \left|\dfrac{\partial \mathbf{r}}{\partial v}\right|$ and $h_3 = \left|\dfrac{\partial \mathbf{r}}{\partial w}\right|$ all are non-zero (in general) and we have vectors

$$\hat{\mathbf{e}}_1 = \frac{1}{h_1}\frac{\partial \mathbf{r}}{\partial u}, \quad \hat{\mathbf{e}}_2 = \frac{1}{h_2}\frac{\partial \mathbf{r}}{\partial v}, \quad \hat{\mathbf{e}}_3 = \frac{1}{h_3}\frac{\partial \mathbf{r}}{\partial w}$$

as the unit vectors in the co-ordinate directions.

Example 7.17 Determine $\nabla\phi$ in general curvilinear co-ordinates.

Solution Let $\nabla\phi = \alpha_1\hat{\mathbf{e}}_1 + \alpha_2\hat{\mathbf{e}}_2 + \alpha_3\hat{\mathbf{e}}_3$ where α_1, α_2 and α_3 are functions to be determined, and $\hat{\mathbf{e}}_1$, $\hat{\mathbf{e}}_2$ and $\hat{\mathbf{e}}_3$ are unit vectors in the u, v and w directions respectively, u, v and w being curvilinear co-ordinates. Expressing the position vector in these curvilinear co-ordinates yields

$$\mathbf{r} = \mathbf{r}(u, v, w)$$

so $$\mathrm{d}\mathbf{r} = \frac{\partial \mathbf{r}}{\partial u}\mathrm{d}u + \frac{\partial \mathbf{r}}{\partial v}\mathrm{d}v + \frac{\partial \mathbf{r}}{\partial w}\mathrm{d}w$$

Now $$\nabla\phi \cdot \mathrm{d}\mathbf{r} = \mathrm{d}\phi = \frac{\partial \phi}{\partial u}\mathrm{d}u + \frac{\partial \phi}{\partial v}\mathrm{d}v + \frac{\partial \phi}{\partial w}\mathrm{d}w$$

but since $$\frac{\partial \mathbf{r}}{\partial u} = h_1\hat{\mathbf{e}}_1,\ \frac{\partial \mathbf{r}}{\partial v} = h_2\hat{\mathbf{e}}_2, \text{ and } \frac{\partial \mathbf{r}}{\partial w} = h_3\hat{\mathbf{e}}_3$$

we also have $$\nabla\phi \cdot \mathrm{d}\mathbf{r} = \alpha_1 h_1 \mathrm{d}u + \alpha_2 h_2 \mathrm{d}v + \alpha_3 h_3 \mathrm{d}w$$

whence $$\alpha_1 h_1 = \frac{\partial \phi}{\partial u},\ \alpha_2 h_2 = \frac{\partial \phi}{\partial v} \text{ and } \alpha_3 h_3 = \frac{\partial \phi}{\partial w}$$

Thus, in curvilinear co-ordinates $\nabla\phi = \dfrac{\hat{\mathbf{e}}_1}{h_1}\dfrac{\partial \phi}{\partial u} + \dfrac{\hat{\mathbf{e}}_2}{h_2}\dfrac{\partial \phi}{\partial v} + \dfrac{\hat{\mathbf{e}}_3}{h_3}\dfrac{\partial \phi}{\partial w}$.

Example 7.18 Find $\nabla \cdot \mathbf{u}$ in general curvilinear co-ordinates.

Solution Using the expression for $\nabla\phi$ derived in the last example, namely $\nabla\phi = \dfrac{\hat{\mathbf{e}}_1}{h_1}\dfrac{\partial \phi}{\partial u} + \dfrac{\hat{\mathbf{e}}_2}{h_2}\dfrac{\partial \phi}{\partial v} + \dfrac{\hat{\mathbf{e}}_3}{h_3}\dfrac{\partial \phi}{\partial w}$, putting $\phi = u$ gives $\nabla u = \dfrac{\hat{\mathbf{e}}_1}{h_1}$. Hence, similarly $\nabla v = \dfrac{\hat{\mathbf{e}}_2}{h_2}$ and $\nabla w = \dfrac{\hat{\mathbf{e}}_3}{h_3}$. This means that $\nabla v \times \nabla w = \dfrac{\hat{\mathbf{e}}_2}{h_2} \times \dfrac{\hat{\mathbf{e}}_3}{h_3} = \dfrac{\hat{\mathbf{e}}_1}{h_2 h_3}$ since $\hat{\mathbf{e}}_1$, $\hat{\mathbf{e}}_2$ and $\hat{\mathbf{e}}_3$ form a right-handed system. Therefore $\hat{\mathbf{e}}_1 = h_2 h_3 \nabla v \times \nabla w$ and hence $\nabla \cdot (h_1\hat{\mathbf{e}}_1) = \nabla \cdot (h_1 h_2 h_3 \nabla v \times \nabla w) =$ $\nabla(A_1 h_2 h_3) \cdot \nabla v \times \nabla w + A_1 h_2 h_3 \nabla \cdot (\nabla v \times \nabla w)$ using the identity $\nabla \cdot (\alpha\mathbf{A}) = \nabla\alpha \cdot \mathbf{A} + \alpha\nabla \cdot \mathbf{A}$. Use of the identity $\nabla \cdot (\mathbf{A} \times \mathbf{B}) = \mathbf{B} \cdot \nabla \times \mathbf{A} - \mathbf{A} \cdot \nabla \times \mathbf{B}$ together with $\nabla \times (\nabla\phi) = \mathbf{0}$ reveals that $\nabla \cdot (\nabla v \times \nabla w) = 0$ and thus

$\nabla \cdot (A_1\hat{\mathbf{e}}_1) = \nabla(A_1 h_2 h_3)\dfrac{\hat{\mathbf{e}}_1}{h_2 h_3}$. Using the curvilinear expression for $\nabla(A_1 h_2 h_3)$ this becomes

$$\begin{aligned}\nabla \cdot (A_1\hat{\mathbf{e}}_1) &= \left[\frac{\hat{\mathbf{e}}_1}{h_1}\frac{\partial}{\partial u}(A_1 h_2 h_3) + \frac{\hat{\mathbf{e}}_2}{h_2}\frac{\partial}{\partial v}(A_1 h_2 h_3) + \frac{\hat{\mathbf{e}}_3}{h_3}\frac{\partial}{\partial w}(A_1 h_2 h_3)\right] \cdot \frac{\hat{\mathbf{e}}_1}{h_2 h_3} \\ &= \frac{1}{h_1 h_2 h_3}\frac{\partial}{\partial u}(A_1 h_2 h_3)\end{aligned}$$

Similarly, $\nabla \cdot (A_2\hat{\mathbf{e}}_2) = \dfrac{1}{h_1 h_2 h_3}\dfrac{\partial}{\partial v}(A_2 h_3 h_1)$ and $\nabla \cdot (A_3\hat{\mathbf{e}}_3) = \dfrac{1}{h_1 h_2 h_3}\dfrac{\partial}{\partial w}(A_3 h_1 h_2)$. The full expression for divergence in curvilinear co-ordinates is thus

$$\nabla \cdot \mathbf{A} = \frac{1}{h_1 h_2 h_3}\left(\frac{\partial}{\partial u}(A_1 h_2 h_3) + \frac{\partial}{\partial v}(A_2 h_3 h_1) + \frac{\partial}{\partial w}(A_3 h_1 h_2)\right)$$

It is worth emphasising here that the components of the arbitrary vector $\mathbf{A}$ on the right-hand side of this expression are its components *in the curvilinear system* and not its components in the everyday sense.

Example 7.19 Express $\nabla \times \mathbf{A}$ in general curvilinear co-ordinates.

Solution Writing $\mathbf{A} = A_1\hat{\mathbf{e}}_1 + A_2\hat{\mathbf{e}}_2 + A_3\hat{\mathbf{e}}_3$ where $\hat{\mathbf{e}}_1$, $\hat{\mathbf{e}}_2$ and $\hat{\mathbf{e}}_3$ are unit vectors in the three co-ordinate directions, and A_1, A_2 and A_3 are the components of $\mathbf{A}$ in these directions, we proceed in a similar fashion to Example 7.18. We have already derived

$$\nabla u = \frac{\hat{\mathbf{e}}_1}{h_1}$$

$$\nabla v = \frac{\hat{\mathbf{e}}_2}{h_2}$$

and $$\nabla w = \frac{\hat{\mathbf{e}}_3}{h_3}$$

$$\text{hence } \nabla \times (A_1\hat{\mathbf{e}}_1) = \nabla \times (A_1h_1\nabla u)$$
$$= \nabla(A_1h_1) \times \nabla u + A_1h_1\nabla \times (\nabla u)$$
$$= \left(\frac{\hat{\mathbf{e}}_1}{h_1}\frac{\partial}{\partial u}(A_1h_1) + \frac{\hat{\mathbf{e}}_2}{h_2}\frac{\partial}{\partial v}(A_1h_1) + \frac{\hat{\mathbf{e}}_3}{h_3}\frac{\partial}{\partial w}(A_1h_1)\right) \times \nabla u + \mathbf{0}$$
$$= -\frac{\hat{\mathbf{e}}_3}{h_1h_2}\frac{\partial}{\partial v}(A_1h_1) + \frac{\hat{\mathbf{e}}_2}{h_1h_3}\frac{\partial}{\partial w}(A_1h_1)$$

using $\nabla u = \dfrac{\hat{\mathbf{e}}_1}{h_1}$, $\hat{\mathbf{e}}_2 \times \hat{\mathbf{e}}_1 = -\hat{\mathbf{e}}_3$ and $\hat{\mathbf{e}}_3 \times \hat{\mathbf{e}}_1 = \hat{\mathbf{e}}_2$. Similarly, for the other components we have

$$\nabla \times (A_2\hat{\mathbf{e}}_2) = \nabla \times (A_2h_2\nabla v) = \frac{\hat{\mathbf{e}}_3}{h_1h_2}\frac{\partial}{\partial u}(A_2h_2) - \frac{\hat{\mathbf{e}}_1}{h_3h_2}\frac{\partial}{\partial w}(A_2h_2) \text{ and}$$
$$\nabla \times (A_3\hat{\mathbf{e}}_3) = \nabla \times (A_3h_3\nabla w) = -\frac{\hat{\mathbf{e}}_2}{h_3h_1}\frac{\partial}{\partial u}(A_3h_3) + \frac{\hat{\mathbf{e}}_1}{h_2h_3}\frac{\partial}{\partial v}(A_3h_3), \text{ whence}$$

$$\nabla \times \mathbf{A} = \nabla \times (A_1\hat{\mathbf{e}}_1 + A_2\hat{\mathbf{e}}_2 + A_3\hat{\mathbf{e}}_3) = \hat{\mathbf{e}}_1\left(\frac{1}{h_2h_3}\frac{\partial}{\partial v}(A_3h_3) - \frac{1}{h_2h_3}\frac{\partial}{\partial w}(A_2h_2)\right) +$$
$$\hat{\mathbf{e}}_2\left(\frac{1}{h_3h_1}\frac{\partial}{\partial w}(A_1h_1) - \frac{1}{h_3h_1}\frac{\partial}{\partial u}(A_3h_3)\right) +$$
$$\hat{\mathbf{e}}_3\left(\frac{1}{h_1h_2}\frac{\partial}{\partial u}(A_2h_2) - \frac{1}{h_1h_2}\frac{\partial}{\partial v}(A_1h_1)\right)$$

$$\text{or } \nabla \times \mathbf{A} = \frac{1}{h_1h_2h_3}\left[h_1\hat{\mathbf{e}}_1\left(\frac{\partial}{\partial v}(A_3h_3) - \frac{\partial}{\partial w}(A_2h_2)\right) + \text{two similar terms}\right]$$
$$= \frac{1}{h_1h_2h_3}\begin{vmatrix} h_1\hat{\mathbf{e}}_1 & h_2\hat{\mathbf{e}}_2 & h_3\hat{\mathbf{e}}_3 \\ \dfrac{\partial}{\partial u} & \dfrac{\partial}{\partial v} & \dfrac{\partial}{\partial w} \\ h_1A_1 & h_2A_2 & h_3A_3 \end{vmatrix} \text{ in determinant form.}$$

Example 7.20

Find expressions for h_1, h_2, h_3, $\hat{\mathbf{e}}_1$, $\hat{\mathbf{e}}_2$ and $\hat{\mathbf{e}}_3$ in cylindrical polar and spherical polar co-ordinates. Hence evaluate the following expressions:
(a) ∇r^n, n is a constant,
(b) $\nabla \times (\mathbf{r}f(r))$ (see Example 7.15),
(c) $\nabla \cdot (r^n\hat{\mathbf{r}})$, n is a constant,
(d) $\nabla \cdot \mathbf{F}$ where $\mathbf{F} = R\sin^2\theta\hat{\theta} - z\mathbf{k}$. Determine the value of θ where $0 < \theta < \frac{1}{2}\pi$ for which $\nabla \cdot \mathbf{F} = 0$.

Solution

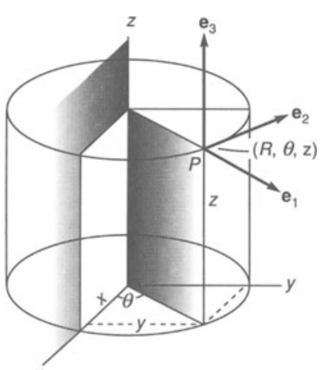


Figure 7.4 Cylindrical polar co-ordinates (R, θ, z). [Adapted from M.R. Spiegel, *Vector Analysis*, Figure 4, page 138, Schaum, McGraw-Hill, New York, 1959]

Cylindrical polar co-ordinates, shown as (R, θ, z) in Figure 7.4 are related to Cartesian co-ordinates (x, y, z) by the equations

$$x = R\cos\theta$$
$$y = R\sin\theta$$
$$z = z$$

From these $\dfrac{\partial\mathbf{r}}{\partial R} = \dfrac{\partial}{\partial R}(\mathbf{i}R\cos\theta + \mathbf{j}R\sin\theta + \mathbf{k}z)$
$= \mathbf{i}\cos\theta + \mathbf{j}\sin\theta$

and $\dfrac{\partial\mathbf{r}}{\partial\theta} = \dfrac{\partial}{\partial\theta}(\mathbf{i}R\cos\theta + \mathbf{j}R\sin\theta + \mathbf{k}z)$
$= -\mathbf{i}R\sin\theta + \mathbf{j}R\cos\theta$

and, trivially $\dfrac{\partial\mathbf{r}}{\partial z} = \mathbf{k}$

Therefore, $\left|\dfrac{\partial\mathbf{r}}{\partial R}\right| = h_1 = 1$, $\left|\dfrac{\partial\mathbf{r}}{\partial\theta}\right| = h_2 = R$, and $\left|\dfrac{\partial\mathbf{r}}{\partial z}\right| = h_3 = 1$. Also,

$$\hat{\mathbf{e}}_1 = \mathbf{i}\cos\theta + \mathbf{j}\sin\theta,\ \hat{\mathbf{e}}_2 = -\mathbf{i}\sin\theta + \mathbf{j}\cos\theta, \text{ and } \hat{\mathbf{e}}_3 = \mathbf{k}$$

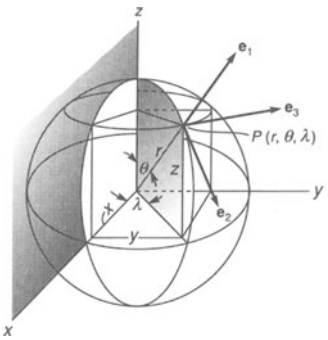


Figure 7.5 Spherical polar co-ordinates (r, θ, λ). [Adapted from M.R. Spiegel, *Vector Analysis,* Figure 5, page 138, Schaum, McGraw-Hill, New York, 1959]

Spherical polar co-ordinates are shown in Figure 7.5.
Writing them as (r, θ, λ), they are related to Cartesian co-ordinates (x, y, z) via the equations

$$x = r\sin\theta\cos\lambda$$
$$y = r\sin\theta\sin\lambda$$
$$z = r\cos\theta$$

whence

$$\frac{\partial \mathbf{r}}{\partial r} = \frac{\partial}{\partial r}(\mathbf{i}r\sin\theta\cos\lambda + \mathbf{j}r\sin\theta\sin\lambda + \mathbf{k}r\cos\theta)$$
$$= \mathbf{i}\sin\theta\cos\lambda + \mathbf{j}\sin\theta\sin\lambda + \mathbf{k}\cos\theta$$
$$\frac{\partial \mathbf{r}}{\partial \theta} = \frac{\partial}{\partial \theta}(\mathbf{i}r\sin\theta\cos\lambda + \mathbf{j}r\sin\theta\sin\lambda + \mathbf{k}r\cos\theta)$$
$$= \mathbf{i}r\cos\theta\cos\lambda + \mathbf{j}r\cos\theta\sin\lambda - \mathbf{k}\sin\theta$$
$$\frac{\partial \mathbf{r}}{\partial \lambda} = \frac{\partial}{\partial \lambda}(\mathbf{i}r\sin\theta\cos\lambda + \mathbf{j}r\sin\theta\sin\lambda + \mathbf{k}r\cos\theta)$$
$$= -\mathbf{i}r\sin\theta\sin\lambda + \mathbf{j}r\sin\theta\cos\lambda$$

Proceeding as with cylindrical polar co-ordinates, we take the modulus of each derivative to find the h's, so $\left|\frac{\partial \mathbf{r}}{\partial r}\right| = h_1 = 1$, $\left|\frac{\partial \mathbf{r}}{\partial \theta}\right| = h_2 = r$ and $\left|\frac{\partial \mathbf{r}}{\partial \lambda}\right| = h_3 = r\sin\theta$. Finally,

$$\hat{\mathbf{e}}_1 = \mathbf{i}\sin\theta\cos\lambda + \mathbf{j}\sin\theta\sin\lambda + \mathbf{k}\cos\theta$$
$$\hat{\mathbf{e}}_2 = \mathbf{i}\cos\theta\cos\lambda + \mathbf{j}\cos\theta\sin\lambda - \mathbf{k}\sin\theta \qquad \text{and}$$
$$\hat{\mathbf{e}}_3 = -\mathbf{i}\sin\lambda + \mathbf{j}\cos\lambda$$

It is easily verified that in each co-ordinate system, $\hat{\mathbf{e}}_1$, $\hat{\mathbf{e}}_2$ and $\hat{\mathbf{e}}_3$ are mutually orthogonal. We are now ready to tackle the problems, the first three of which are in spherical polar co-ordinates:

(a) $\nabla r^n = \hat{\mathbf{r}}\ \frac{\partial}{\partial r}(r^n)$ since there is no θ or λ dependence
$= n\hat{\mathbf{r}}r^{n-1} = n\mathbf{r}r^{n-2}\ (\mathbf{r} = \hat{\mathbf{r}}r)$

(b) $\nabla \times (\mathbf{r}f(r)) = \begin{vmatrix} \hat{\mathbf{r}} & r\hat{\boldsymbol{\theta}} & r\sin\theta\hat{\boldsymbol{\lambda}} \\ \frac{\partial}{\partial r} & \frac{\partial}{\partial \theta} & \frac{\partial}{\partial \lambda} \\ f(r) & 0 & 0 \end{vmatrix} = \mathbf{0}$, this confirms the result of Example 7.15 (rather more briefly!)

(c) $\nabla \cdot (r^n\hat{\mathbf{r}}) = \frac{1}{r^2}\frac{\partial}{\partial r}(r^2r^n) = \frac{1}{r^2}(n + 2)(r^{n+1}) = (n + 2)r^{n+1}$ since again there is no dependence on θ or λ.

(d) For this last exercise, we use cylindrical polar co-ordinates.
In cylindrical polars, $\nabla \cdot \mathbf{F} = \frac{1}{R}\left(\frac{\partial}{\partial R}(RF_1) + \frac{\partial}{\partial \theta}(F_2) + \frac{\partial}{\partial z}(RF_3)\right)$

$$\nabla(R\sin^2\theta\hat{\boldsymbol{\theta}} - z\mathbf{k}) = \frac{1}{R}\left(\frac{\partial}{\partial \theta}(R\sin^2\theta) - \frac{\partial}{\partial z}(Rz)\right)$$
$$= \frac{1}{R}(2R\sin\theta\cos\theta - R) = \sin 2\theta - 1$$

This expression is zero when $\sin 2\theta = 1$, that is $\theta = \pi/4$ in $0 \leq \theta \leq \pi/2$.

Example 7.21 Express $\nabla^2\phi$ in general curvilinear co-ordinates. Hence write down expressions for $\nabla^2\phi$ in cylindrical and spherical polar co-ordinates. From these expressions solve $\nabla^2\phi = 0$ if
(a) ϕ depends only on R in cylindrical polar co-ordinates,
(b) ϕ depends only on r in spherical polar co-ordinates.

Solution The Laplacian $\nabla^2\phi$ is defined as $\nabla \cdot (\nabla\phi)$.
Using $\nabla\phi = \frac{\hat{\mathbf{e}}_1}{h_1}\frac{\partial \phi}{\partial u} + \frac{\hat{\mathbf{e}}_2}{h_2}\frac{\partial \phi}{\partial v} + \frac{\hat{\mathbf{e}}_3}{h_3}\frac{\partial \phi}{\partial w}$ and

$$\nabla \cdot \mathbf{A} = \frac{1}{h_1h_2h_3}\left(\frac{\partial}{\partial u}(A_1h_2h_3) + \frac{\partial}{\partial v}(A_2h_3h_1) + \frac{\partial}{\partial w}(A_3h_1h_2)\right) \text{ we have}$$

$$\nabla^2\phi = \frac{1}{h_1h_2h_3}\left(\frac{\partial}{\partial u}\left(\frac{h_2h_3}{h_1}\frac{\partial\phi}{\partial u}\right) + \frac{\partial}{\partial v}\left(\frac{h_3h_1}{h_2}\frac{\partial\phi}{\partial v}\right) + \frac{\partial}{\partial w}\left(\frac{h_1h_2}{h_3}\frac{\partial\phi}{\partial w}\right)\right)$$

(a) For cylindrical polar co-ordinates, $h_1 = 1$, $h_2 = R$, $h_3 = 1$. Thus in cylindrical polars (R, θ, z) we have

$$\begin{aligned}\nabla^2\phi &= \frac{1}{R}\left(\frac{\partial}{\partial R}\left(R\frac{\partial\phi}{\partial R}\right) + \frac{1}{R}\frac{\partial}{\partial\theta}\left(\frac{1}{R}\frac{\partial\phi}{\partial\theta}\right) + \frac{\partial}{\partial z}\left(R\frac{\partial\phi}{\partial z}\right)\right)\\ &= \frac{1}{R}\frac{\partial}{\partial R}\left(R\frac{\partial\phi}{\partial R}\right) + \frac{1}{R^2}\frac{\partial^2\phi}{\partial\theta^2} + \frac{\partial^2\phi}{\partial z^2}\end{aligned}$$

which, if ϕ depends only on R, contains only the first term. Thus $\nabla^2\phi = \frac{1}{R}\frac{\mathrm{d}}{\mathrm{d}R}\left(R\frac{\mathrm{d}\phi}{\mathrm{d}R}\right) = 0$ which integrates once to $R\frac{\mathrm{d}\phi}{\mathrm{d}R} = A_1$ where A_1 is an arbitrary constant. Rearranging this as $\frac{\mathrm{d}\phi}{\mathrm{d}R} = \frac{A_1}{R}$ and integrating again gives $\phi = A_1\ln R + B_1$, where B_1 is a second arbitrary constant.

(b) In spherical polars (r, θ, λ) $h_1 = 1$, $h_2 = r$, $h_3 = r\sin\theta$ so

$$\begin{aligned}\nabla^2\phi &= \frac{1}{r^2\sin\theta}\left(\frac{\partial}{\partial r}\left(\frac{r \cdot r\sin\theta}{1}\frac{\partial\phi}{\partial r}\right) + \frac{\partial}{\partial\theta}\left(\frac{r\sin\theta \cdot 1}{r}\frac{\partial\phi}{\partial\theta}\right) + \frac{\partial}{\partial\lambda}\left(\frac{1 \cdot r}{r\sin\theta}\frac{\partial\phi}{\partial\lambda}\right)\right)\\ &= \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\phi}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\phi}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\phi}{\partial\lambda^2}\end{aligned}$$

If ϕ depends only on r then this reduces to $\nabla^2\phi = \frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}\phi}{\mathrm{d}r}\right)$ which when equated to zero yields $\frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}\phi}{\mathrm{d}r}\right) = 0$ or $\frac{\mathrm{d}\phi}{\mathrm{d}r} = \frac{A_2}{r^2}$ on intergrating once, where A_2 is an arbitrary constant. Integrating again finally gives

$$\phi = B_2 - \frac{A_2}{r}$$

where B_2 is a second arbitrary constant.

Example 7.22 If $\mathbf{F} = (R\sin\theta\cos\theta + z\cos\theta)\hat{\mathbf{R}} + (R\cos^2\theta - z\sin\theta)\hat{\boldsymbol{\theta}} + R\sin\theta\hat{\mathbf{k}}$ use cylindrical polar co-ordinates to show that $\nabla \times (\nabla \times \mathbf{F}) = \mathbf{0}$.

Solution This is a purely mechanical calculation. It is a good exercise in using the formula for curl in cylindrical polar co-ordinates. (Some will say that this type of exercise has been or soon will be, superseded by computer algebra and should therefore be consigned to the dustbin of history alongside long division by hand. Such debates are for another forum!)

$$\nabla \times \mathbf{F} = \frac{1}{R}\begin{vmatrix} \hat{\mathbf{R}} & R\hat{\boldsymbol{\theta}} & \hat{\mathbf{k}} \\ \frac{\partial}{\partial R} & \frac{\partial}{\partial\theta} & \frac{\partial}{\partial z} \\ R\sin\theta\cos\theta + z\cos\theta & R^2\cos^2\theta \equiv R\sin\theta & R\sin\theta \end{vmatrix}$$

which is $\qquad = \hat{\mathbf{R}}(\cos\theta + \sin\theta) + \hat{\boldsymbol{\theta}}(\cos\theta - \sin\theta) + \hat{\mathbf{k}}$

where we have used the trigonometric identities $\sin\theta\cos\theta = \frac{1}{2}\sin2\theta$ and $\cos2\theta = 2\cos^2\theta - 1$. So therefore

$$\nabla \times (\nabla \times \mathbf{F}) = \frac{1}{R}\begin{vmatrix} \hat{\mathbf{R}} & R\hat{\boldsymbol{\theta}} & \hat{\mathbf{k}} \\ \frac{\partial}{\partial R} & \frac{\partial}{\partial\theta} & \frac{\partial}{\partial z} \\ \cos\theta + \sin\theta & R(\cos\theta - \sin\theta) & 1 \end{vmatrix}$$

$$\begin{aligned}&= \frac{1}{R}[0\hat{\mathbf{R}} + R\hat{\boldsymbol{\theta}}(0 - 0) + \hat{\mathbf{k}}(\cos\theta - \sin\theta - (-\sin\theta + \cos\theta))]\\ &= \mathbf{0} \text{ as required.}\end{aligned}$$

7.3 Exercises

7.1. Find the unit normal to the surface $x^2y^2 + y^2z^2 + z^2x^2 = k$ (where k is a constant) at the points $(1, 1, 1)$ and $(1, 0, -1)$. Show that the normal at the point $(a, 0, -a)$ is perpendicular to the normal at the point $(a, 0, a)$ for all values of the constant a.

7.2. (a) If $\phi = x^n + y^n + z^n$ show that $\mathbf{r} \cdot \nabla\phi = n\phi$.
(b) If $\phi = x^a y^b z^c$ show that $\mathbf{r} \cdot \nabla\phi = (a + b + c)\phi$.
In this question, n, a, b and c are all constants.

7.3. Find the direction cosines of the normal to the ellipsoid $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 3$ at the point (a, b, c).

7.4. Establish that for the twice differentiable scalar functions ϕ and ψ

$$\nabla\left(\frac{\phi}{\psi}\right) = \frac{\psi\nabla\phi - \phi\nabla\psi}{\psi^2}.$$

7.5. Determine the following scalar functions ϕ to within an arbitrary constant given that $\mathbf{F} = \nabla\phi$:

(a) $\mathbf{F} = \frac{1}{x}\mathbf{i} + \frac{1}{y}\mathbf{j} + \frac{1}{z}\mathbf{k}$,
(b) $\mathbf{F} = yz\cos xyz\mathbf{i} + zx\cos xyz\mathbf{j} + xy\cos xyz\mathbf{k}$,
(c) $\mathbf{F} = 2xy^3z^4\mathbf{i} + 3x^2y^2z^4\mathbf{j} + 4x^2y^3z^3\mathbf{k}$,
(d) $\mathbf{F} = ((y + xy^2z)\mathbf{i} + (x + x^2yz)\mathbf{j} + \frac{1}{2}x^2y^2\mathbf{k})e^{xyz}$.

7.6. If $\mathbf{E} = -\nabla\phi$ and $\phi = \frac{\mathbf{p}\cdot\mathbf{r}}{r^3}$, determine $\mathbf{E}$.

7.7. Find the directional derivatives of the following functions in the directions indicated:
(a) $\phi = x^2z + 2xy^2 - yz^2$ along $2\mathbf{i} + 3\mathbf{j} - 4\mathbf{k}$ at $(1, 2, -1)$,
(b) $\phi = x^2y + y^2z + z^2x$ along $4\mathbf{i} + 2\mathbf{j} - 5\mathbf{k}$ at $(1, -1, 2)$,
(c) $\phi = (x + 3y)^2 + (2y - z)^2$ along $4\mathbf{i} + 2\mathbf{j} - 5\mathbf{k}$ at $(1, 1, 0)$.

7.8. Find the direction from the origin along which the temperature field $T(x, y, z) = T_0(1 + cz + by)e^{ax}$ changes most rapidly.

7.9. Find the unit normal to the surface $xz^2 + x^2y = z - 1$ at the point $(1, -3, 2)$ and hence find the tangent plane.

7.10. Find the acute angle between the surfaces $xy^2z = 3x + z^2$ and $3x^2 - y^2 + 2z = 1$ at the point $(1, -2, 1)$.

7.11. If $\mathbf{F}$ is the vector field $xy\mathbf{i} + yz\mathbf{j}$ find $\nabla \cdot \mathbf{F}$, $\nabla \times \mathbf{F}$ and $\nabla(\nabla \cdot \mathbf{F})$.

7.12. Find $\nabla \cdot \mathbf{A}$ for the following fields:
(a) $\mathbf{A} = 2x^2z\mathbf{i} - xy^2z\mathbf{j} + 3yz^2\mathbf{k}$,
(b) $\mathbf{A} = x^2y\mathbf{i} - xyz\mathbf{j} + yz^2\mathbf{k}$,
(c) $\mathbf{A} = \mathbf{a}e^{xyz}$ where $\mathbf{a}$ is a constant vector.

7.13. Find $\nabla \times \mathbf{A}$ for the following fields:
(a) $\mathbf{A} = xz^3\mathbf{i} - 2x^2yz\mathbf{j} + 2yz^4\mathbf{k}$,
(b) $\mathbf{A} = \phi\nabla\phi$, where ϕ is any scalar differentiable function,
(c) $\mathbf{A} = \mathbf{r}/r$, where $\mathbf{r}$ is the position vector.

7.14. Establish the following vector identities:
(a) $\nabla \times (\phi\mathbf{A}) = \nabla\phi \times \mathbf{A} + \phi\nabla \times \mathbf{A}$,
(b) $\nabla(\mathbf{A} \cdot \mathbf{B}) = (\mathbf{B} \cdot \nabla)\mathbf{A} + (\mathbf{A} \cdot \nabla)\mathbf{B} + \mathbf{B} \times (\nabla \times \mathbf{A}) + \mathbf{A} \times (\nabla \times \mathbf{B})$,
(c) $\nabla \times (\mathbf{A} \times \mathbf{B}) = (\mathbf{B} \cdot \nabla)\mathbf{A} - (\mathbf{A} \cdot \nabla)\mathbf{B} - \mathbf{B}(\nabla \cdot \mathbf{A}) + \mathbf{A}(\nabla \cdot \mathbf{B})$,

where ϕ is a differentiable scalar function, and $\mathbf{A}$ and $\mathbf{B}$ are differentiable vector functions, by considering the components of $\mathbf{A}$ and $\mathbf{B}$. Hence deduce that $(\mathbf{u} \cdot \nabla)\mathbf{u} = \frac{1}{2}\nabla(\mathbf{u}^2) - \mathbf{u} \times (\nabla \times \mathbf{u})$. What is a potential function for $(\mathbf{u} \cdot \nabla)\mathbf{u}$ if $\mathbf{u}$ is a fluid velocity and the flow is irrotational?

7.15. Show that the field $\nabla\phi \times \nabla\psi$ is equal to $\nabla \times \mathbf{A}$ for some differentiable vector field $\mathbf{A}$. (*Hint*: show that $\nabla\phi \times \nabla\psi$ is solenoidal.)

7.16. Show that the scalar triple product

$$\nabla\phi \cdot \nabla\psi \times \nabla\Omega = \frac{\partial(\phi, \psi, \Omega)}{\partial(x, y, z)}.$$

Deduce a necessary and sufficient condition for ϕ, ψ and Ω to be functionally related.

7.17. Prove that $\nabla^2(\phi\psi) = \phi\nabla^2\psi + 2\nabla\phi \cdot \nabla\psi + \psi\nabla^2\phi$ for two twice differentiable scalar functions ϕ and ψ.

7.18. Maxwell's equations can take the form

$$\nabla \times \mathbf{H} = \frac{1}{c}\frac{\partial\mathbf{E}}{\partial t}, \qquad \nabla \cdot \mathbf{H} = 0,$$

$$\nabla \times \mathbf{E} = -\frac{1}{c}\frac{\partial\mathbf{H}}{\partial t}, \qquad \nabla \cdot \mathbf{E} = 4\pi\rho,$$

where $\mathbf{H}$ and $\mathbf{E}$ are magnetic and electric fields respectively, $\rho = \rho(x, y, z)$ is a scalar field, and c is the (constant) speed of light. Show that it is possible to write a solution to these equations in the form

$$\mathbf{E} = -\nabla\phi - \frac{1}{c}\frac{\partial\mathbf{A}}{\partial t}, \qquad \mathbf{H} = \nabla \times \mathbf{A}$$

where $\mathbf{A}$ and ϕ, called the vector and scalar potentials respectively, satisfy the three equations

$$\nabla \cdot \mathbf{A} + \frac{1}{c}\frac{\partial\phi}{\partial t} = 0$$

$$\nabla^2\phi - \frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2} = -4\pi\rho$$

$$\nabla^2\mathbf{A} - \frac{1}{c^2}\frac{\partial^2\mathbf{A}}{\partial t^2} = \mathbf{0}$$

7.19. Show that Maxwell's equations in free space

$$\nabla \times \mathbf{H} = \frac{1}{c}\frac{\partial\mathbf{E}}{\partial t}, \qquad \nabla \cdot \mathbf{H} = 0,$$

$$\nabla \times \mathbf{E} = -\frac{1}{c}\frac{\partial\mathbf{H}}{\partial t}, \qquad \nabla \cdot \mathbf{E} = 0,$$

are satisfied by

$$\mathbf{H} = \frac{1}{c}\nabla \times \left(\frac{\partial\mathbf{Z}}{\partial t}\right)$$

and

$$\mathbf{E} = \nabla \times (\nabla \times \mathbf{Z})$$

provided the vector field $\mathbf{Z}$ (sometimes called the Hertzian vector) obeys the wave equation

$$\nabla^2\mathbf{Z} = \frac{1}{c^2}\frac{\partial^2\mathbf{Z}}{\partial t^2}$$

7.20. If the velocity of a fluid at the point (x, y, z) takes the general linear but two-dimensional form

$$\mathbf{u} = (ax + by)\mathbf{i} + (kx + ly)\mathbf{j}$$

find the conditions on the constants a, b, k and l such that the mass is conserved ($\nabla \cdot \mathbf{u} = 0$), and the flow is irrotational ($\nabla \times \mathbf{u} = \mathbf{0}$). Verify in this case that $\mathbf{u}$ takes the form

$$\mathbf{u} = \nabla[\tfrac{1}{2}(ax^2 + 2bxy - ay^2)]$$

7.21. Defining toroidal curvilinear co-ordinates (ρ, α, β) through the transformation equations:

$$x = (a - \rho\cos\alpha)\cos\beta,$$
$$y = (a - \rho\cos\alpha)\sin\beta, \text{ and}$$
$$z = \rho\sin\alpha$$

where a is a constant such that $\rho < a$, find the quantities $\hat{\mathbf{e}}_1$, $\hat{\mathbf{e}}_2$, $\hat{\mathbf{e}}_3$ and show that this is an orthogonal system.
Show further that

$$h_1 = 1, \quad h_2 = \rho, \quad h_3 = a - \rho\cos\alpha$$

7.22. From the expressions for $\nabla \cdot \mathbf{A}$ and $\nabla \times \mathbf{A}$ derived in Examples 7.17 and 7.18 show that $\nabla \cdot (\nabla \times \mathbf{A}) = 0$ and $\nabla \times (\nabla\phi) = \mathbf{0}$ in general curvilinear co-ordinates.

7.23. Using the toroidal system of co-ordinates (ρ, α, β) introduced in Exercise 7.21, write down the expression $\nabla \cdot (\phi\hat{\mathbf{e}}_\rho)$. If $\phi = \phi(\rho, \alpha)$, and $\nabla \cdot (\phi\hat{\mathbf{e}}_\rho) = 0$ find the general form of $\phi = \phi(\rho, \alpha)$ given that $\phi(a, \alpha) = (\sin^2\tfrac{1}{2}\alpha)^{-1}$.

Topic Guide

Definition
Parameterisation
Closed Contours
Conservative Fields

8 Line Integrals

8.1 Fact Sheet

Let $\mathbf{F}(t) = \mathbf{i}F_1(t) + \mathbf{j}F_2(t) + \mathbf{k}F_3(t)$ be a vector function dependent on a single scalar variable t. Suppose also that $\mathbf{r}(t) = \mathbf{i}x(t) + \mathbf{j}y(t) + \mathbf{k}z(t)$ represents a curve (see Chapter 6) then the *line integral*

$$\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r}$$

along the curve C is defined by

$$\int_{t_1}^{t_2} \left(F_1 \frac{\mathrm{d}x}{\mathrm{d}t} + F_2 \frac{\mathrm{d}y}{\mathrm{d}t} + F_3 \frac{\mathrm{d}z}{\mathrm{d}t} \right) \mathrm{d}t$$

where C is the curve described by $\mathbf{r}(t)$ starting at $t = t_1$ and finishing at $t = t_2$. If $t_1 = t_2$ then the curve C is closed and the integral around C is sometimes written

$$\oint_C \mathbf{F} \cdot \mathrm{d}\mathbf{r}$$

Other line or *contour* integrals are defined similarly, namely

$$\int_C \mathbf{F} \times \mathrm{d}\mathbf{r} = \int_{t_1}^{t_2} \mathbf{F} \times \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}\, \mathrm{d}t$$

$$\int_C \phi \mathrm{d}\mathbf{r} = \int_{t_1}^{t_2} \phi \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}\, dt$$

where the integrals on the right-hand side are vectors which are evaluated in the normal way.

If $\mathbf{F}$ is the vector representative of a force and C is any curve (path) in the domain of the force, then the integral $\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r}$ represents the *work done* by the force $\mathbf{F}$ in moving a unit mass along the curve C. If $\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r}$ is independent of the path C then $\mathbf{F}$ is a *conservative* force. For conservative forces, $\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} = 0$ if C is closed. This is equivalent to the condition that $\nabla \times \mathbf{F} = \mathbf{0}$, that is there exists a scalar potential ϕ such that $\mathbf{F} = \nabla\phi$. ϕ is called the *scalar potential* of $\mathbf{F}$.

8.2 Worked Examples

Example 8.1 Evaluate the (scalar) line integral

$$\int_C x^2\, y\mathrm{d}s$$

where C is the semi-circle $x^2 + y^2 = a^2$, $z = 0$, $y \geq 0$.

Solution First of all, let us parameterise the circle as follows:

$$x = a\cos\theta, \qquad y = a\sin\theta \qquad 0 \leq \theta \leq \pi$$

We can do this since, on C, $x^2 + y^2 = a^2(\cos^2\theta + \sin^2\theta) = a^2$ as is required. It is very important, not to say crucial, to parameterise the curve C. Otherwise it is normally impossible to do the line integral. Fortunately, parameterisations are either quite straightforward as in this case, or they are given. There is often more than one way to parameterise a curve, but *any* parameterisation will suffice to evaluate the line integral. The next step is to put all quantities in the body of the integral (the integrand) in terms of the parameter θ.

$$\begin{aligned}\text{Now } (\mathrm{d}s)^2 &= (\mathrm{d}x)^2 + (\mathrm{d}y)^2 + (\mathrm{d}z)^2\\ &= a^2\sin^2\theta(\mathrm{d}\theta)^2 + a^2\cos^2\theta(\mathrm{d}\theta)^2 + (0)^2\\ &= a^2(\mathrm{d}\theta)^2\end{aligned}$$

so that $\mathrm{d}s = a\mathrm{d}\theta$ (this could have been deduced immediately since $a\mathrm{d}\theta$ is the element of arc length of a circle of radius a).

On C, $x^2y = a^3\cos^2\theta\sin\theta$ and so

$$\begin{aligned}\int_C x^2\, y\mathrm{d}s &= \int_0^\pi a^3\cos^2\theta\sin\theta \cdot a\mathrm{d}\theta\\ &= a^4\int_0^\pi \cos^2\theta\sin\theta\mathrm{d}\theta\\ &= a^4\left[-\tfrac{1}{3}\cos^3\theta\right]_0^\pi\\ &= \tfrac{2}{3}a^4\end{aligned}$$

the required answer.

Example 8.2 Evaluate the integral

$$\int_C \mathbf{F}\cdot\mathbf{dr}$$

where C is the semi-circle and $x^2 + y^2 = a^2$, $z = 0$, $y \geq 0$, and $\mathbf{F} = xy(\mathbf{i} + \mathbf{j})$.

Solution Again we emphasise that the first step is to parameterise the curve C. It is the same curve as in the last example, therefore we use the same parameterisation, namely

$x = a\cos\theta$, $y = a\sin\theta$ $\quad 0 \leq \theta \leq \pi$. Since $\mathbf{r} = x\mathbf{i} + y\mathbf{j}$ we have that

on C $\qquad \mathbf{dr} = \mathrm{d}x\mathbf{i} + \mathrm{d}y\mathbf{j} = -a\sin\theta\mathrm{d}\theta\mathbf{i} + a\cos\theta\mathrm{d}\theta\mathbf{j}$

also, on C $\qquad \mathbf{F} = xy(\mathbf{i} + \mathbf{j}) = a^2\sin\theta\cos\theta(\mathbf{i} + \mathbf{j})$

Hence, still on C of course, $\mathbf{F}\cdot\mathbf{dr} = a^3(-\sin^2\theta\cos\theta + \cos^2\theta\sin\theta)\mathrm{d}\theta$ so that

$$\begin{aligned}\int_C \mathbf{F}\cdot\mathbf{dr} &= \int_0^\pi a^3(-\sin^2\theta\cos\theta + \cos^2\theta\sin\theta)\mathrm{d}\theta\\ &= a^3\left[-\tfrac{1}{3}\sin^3\theta - \tfrac{1}{3}\cos^3\theta\right]_0^\pi = \tfrac{2}{3}a^3\end{aligned}$$

Example 8.3 A vector field $\mathbf{F}(x, y, z)$ is given by the expression $\mathbf{F} = 2xyz\mathbf{i} + x^2z\mathbf{j} + x^2y\mathbf{k}$.

(a) Evaluate $\int_C \mathbf{F}\cdot\mathbf{dr}$ between the points (0, 0, 0) and (2, 4, 6) along the two different paths:

(i) $x = u, \quad y = u^2, \quad z = 3u$

(ii) straight lines (0, 0, 0) to (2, 0, 0), (2, 0, 0) to (2, 4, 0) and finally (2, 4, 0) to (2, 4, 6).

(b) Show that $\mathbf{F}$ is a conservative field and deduce that for conservative fields in general, $\int_C \mathbf{F}\cdot\mathbf{dr}$ is independent of the path and thus must be zero for any closed curve.

Solution (a) (i) First of all, we need to find the values of the parameter u that correspond to the points (0,0,0) and (2,4,6). (In this problem, the parameterisation has been prescribed.) It is quickly seen that the values are $u = 0$ and $u = 2$ respectively.

$$\text{On } C \qquad \mathbf{F} = 2u \cdot u^2 \cdot 3u\mathbf{i} + u^2 \cdot 3u\mathbf{j} + u^2 \cdot u^2\mathbf{k}$$
$$= 6u^4\mathbf{i} + 3u^3\mathbf{j} + u^4\mathbf{k}$$
$$\text{and} \qquad \mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$$
$$= u\mathbf{i} + u^2\mathbf{j} + 3u\mathbf{k}$$
$$\text{so that} \qquad \mathbf{dr} = \mathrm{d}u\mathbf{i} + 2u\mathrm{d}u\mathbf{j} + 3\mathrm{d}u\mathbf{k}$$
$$\text{This gives} \qquad \mathbf{F} \cdot \mathbf{dr} = 6u^4\mathrm{d}u + 3u^3 \cdot 2u\mathrm{d}u + u^4 \cdot 3\mathrm{d}u = 15u^4\mathrm{d}u$$
$$\text{Hence} \qquad \int_C \mathbf{F} \cdot \mathbf{dr} = \int_0^2 15u^4\mathrm{d}u = \left[3u^5\right]_0^2 = 3 \cdot 2^5 = 96$$

(ii) For this next part, a diagram is useful. The path between (0, 0, 0) to (2, 4, 6) is a sequence of three straight lines is shown in Figure 8.1.

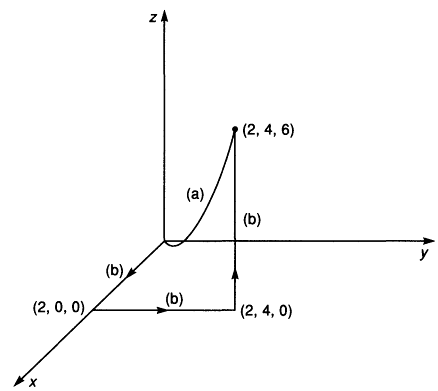


Figure 8.1 The integration paths (a), and the three straight paths (b).

From (0,0,0) to (2,0,0), y and z are identically zero, hence $\mathbf{F}$ ($= 2xyz\mathbf{i} + x^2z\mathbf{j} + x^2y\mathbf{k}$) is also identically equal to zero and so must be $\int_C \mathbf{F} \cdot \mathbf{dr}$ along this line. From (2, 0, 0) to (2, 4, 0), z remains zero, $x = 2$ and y varies from 0 to 4. Hence $\mathbf{F} = 4y\mathbf{k}$. However, $\mathbf{dr} = \mathbf{j}\mathrm{d}y$ (since only y is varying), so $\mathbf{F} \cdot \mathbf{dr} = 4y\mathrm{d}y\mathbf{j} \cdot \mathbf{k} = 0$ once more. Therefore $\int_C \mathbf{F} \cdot \mathbf{dr} = 0$ again along this line. Finally, from (2, 4, 0) to (2, 4, 6) $x = 2$ and $y = 4$ but z varies from 0 to 6 to give

$$\mathbf{F} = 16z\mathbf{i} + 4z\mathbf{j} + 16\mathbf{k}$$
$$\text{and} \qquad \mathbf{dr} = \mathbf{k}\mathrm{d}z \text{ (only } z \text{ varies)}$$
$$\text{Thus} \qquad \mathbf{F} \cdot \mathbf{dr} = 16\mathrm{d}z \text{ and}$$
$$\int_C \mathbf{F} \cdot \mathbf{dr} = \int_0^6 16\mathrm{d}z = 96$$

This is the same result as before. It does not, of course show that $\int_C \mathbf{F} \cdot \mathbf{dr}$ between (0, 0, 0) and (2, 4, 6) is *always* 96 but it at least does not lead us to dismiss this possibility.

(b) In order to show conclusively that $\mathbf{F}$ is conservative, we need to show that $\nabla \times \mathbf{F} = \mathbf{0}$. Using the definition of curl

$$\nabla \times \mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ 2xyz & x^2z & x^2y \end{vmatrix}$$
$$= \mathbf{i}(x^2 - x^2) + \mathbf{j}(2xy - 2xy) + \mathbf{k}(2xz - 2xz)$$
$$= \mathbf{0}$$

This shows that $\mathbf{F}$ is conservative. If $\nabla \times \mathbf{F} = \mathbf{0}$, then there exists a scalar potential $\phi(x, y, z)$ such that $\mathbf{F} = \nabla\phi$ (see Example 7.13). It is the existence of this scalar potential that enables us to prove the last part of the question.

Since $\mathbf{F} = \nabla\phi$, then

$$\mathbf{F} \cdot \mathbf{dr} = \nabla\phi \cdot \mathbf{dr} = \frac{\partial\phi}{\partial x}\mathrm{d}x + \frac{\partial\phi}{\partial y}\mathrm{d}y + \frac{\partial\phi}{\partial z}\mathrm{d}z = \mathrm{d}\phi \text{ (using the chain rule, see Chapter 2).}$$

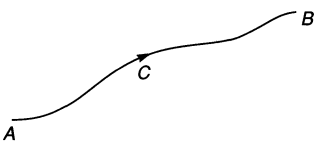


Figure 8.2 An arbitrary path C between two points A and B.

Hence $\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \int_C \mathrm{d}\phi = [\phi]_C$ which represents the change in ϕ as C is traversed. If A is the starting point of C, and B is its end point (see Figure 8.2) then $\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \phi_A - \phi_B$. If A and B are the same point (that is, C is a closed curve) then $\oint_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} = 0$ where the symbol $\oint$ has been used to indicate a closed (as opposed to open) curve.

For this example, $\phi(x, y, z) = x^2 yz + k$, so $\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} = (2, 4, 6) - \phi(0, 0, 0) = 2^2 \times 4 \times 6 = 96$ for any path C connecting the point (0, 0, 0) to the point (2, 4, 6).

Example 8.4 Evaluate the integral $\int_C \mathbf{F} \times \mathrm{d}\mathbf{r}$ where $\mathbf{F} = x\mathbf{i} + 2y\mathbf{j} + 3z\mathbf{k}$ and C is the straight line connecting (0, 0, 0) and (1, 1, 1).

Solution On the straight line C connecting (0, 0, 0) to (1, 1, 1), $x = y = z = t$ where t varies from 0 to 1. Hence, on C

$$\mathbf{F} = t\mathbf{i} + 2t\mathbf{j} + 3t\mathbf{k}$$
$$\mathrm{d}\mathbf{r} = \mathbf{i}\mathrm{d}t + \mathbf{j}\mathrm{d}t + \mathbf{k}\mathrm{d}t$$

whence
$$\mathbf{F} \times \mathrm{d}\mathbf{r} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ t & 2t & 3t \\ \mathrm{d}t & \mathrm{d}t & \mathrm{d}t \end{vmatrix} = (-t\mathbf{i} + 2t\mathbf{j} - t\mathbf{k})\mathrm{d}t$$

and so
$$\begin{aligned} \int_C \mathbf{F} \times \mathrm{d}\mathbf{r} &= \int_0^1 (-t\mathbf{i} + 2t\mathbf{j} - t\mathbf{k})\mathrm{d}t \\ &= \left[-\tfrac{1}{2}t^2\mathbf{i} + t^2\mathbf{j} - \tfrac{1}{2}t^2\mathbf{k}\right]_0^1 \\ &= -\tfrac{1}{2}\mathbf{i} + \mathbf{j} - \tfrac{1}{2}\mathbf{k} \end{aligned}$$

Note that a vector answer is expected since $\int_C \mathbf{F} \times \mathrm{d}\mathbf{r}$ is a vector.

Example 8.5 The formula $\mathbf{H} = I\int_C \dfrac{\mathrm{d}\mathbf{s} \times \mathbf{r}}{r^3}$ can be related to the magnetic field at a point in space with position vector $\mathbf{r}$, due to a current I passing along a perfectly conducting wire of shape C with directed element of arc length $\mathrm{d}\mathbf{s}$ (the Biot–Savart Law). Evaluate $\mathbf{H}$ for the cases
(a) C is the circle $x^2 + y^2 = a^2$,
(b) C is an infinitely long wire along the x-axis.

Solution The reader familiar with electromagnetic theory will be able to spot deficiencies if it is attempted to relate this $\mathbf{H}$ directly to magnetic induction. This is a learning example to see the kind of integrals met in electromagnetic theory. For the real thing, see the next example.
(a) The most natural (and by far the easiest) way of evaluating the contour integral is to use cylindrical polar co-ordinates (R, θ, z) for which

$$\mathbf{r} = R\hat{\mathbf{R}} + z\hat{\mathbf{k}} \text{ (see Figure 8.3)}$$

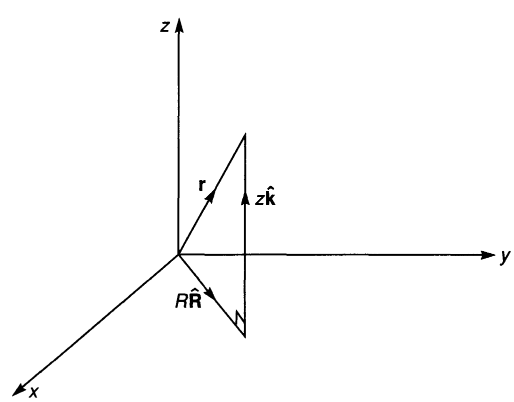


Figure 8.3 $\mathbf{r}$ in cylindrical co-ordinates showing that $\mathbf{r} = R\hat{\mathbf{R}} + z\hat{\mathbf{k}}$.

For the circle $x^2 + y^2 = a^2$, or $R = a$ we have

$$\mathrm{d}\mathbf{s} = a\mathrm{d}\theta\hat{\boldsymbol{\theta}} \text{ (see Figure 8.4)}$$

Hence
$$\mathrm{d}\mathbf{s} \times \mathbf{r} = a\mathrm{d}\theta\hat{\boldsymbol{\theta}} \times (R\hat{\mathbf{R}} + z\hat{\mathbf{k}})$$
$$= -aR\mathrm{d}\theta\hat{\mathbf{k}} + az\mathrm{d}\theta\hat{\mathbf{R}}$$
$((\hat{\mathbf{R}}, \theta, \hat{\mathbf{k}})$ is a right-handed system)

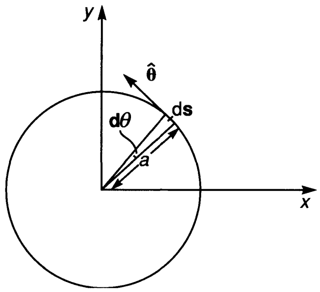


Figure 8.4 Showing that $\mathrm{d}\mathbf{s} = a\mathrm{d}\theta\hat{\boldsymbol{\theta}}$.

and thus

$$\int_C \frac{\mathrm{d}\mathbf{s} \times \mathbf{r}}{r^3} = \int_0^{2\pi} \frac{-aR\hat{\mathbf{k}} + az\hat{\mathbf{R}}}{(R^2 + a^2)^{\frac{3}{2}}}\,\mathrm{d}\theta$$

and now, since $\hat{\mathbf{R}} = \mathbf{i}\cos\theta + \mathbf{j}\sin\theta$, the $\hat{\mathbf{R}}$ term integrates to zero. Thus we obtain
$\displaystyle\int_C \frac{\mathrm{d}\mathbf{s} \times \mathbf{r}}{r^3} = \frac{2\pi a}{(R^2 + a^2)^{\frac{3}{2}}}(-R\hat{\mathbf{k}})$ since there is no dependence on θ.

Thus, $\mathbf{H} = -\dfrac{2\pi aIR\hat{\mathbf{k}}}{(R^2 + a^2)^{\frac{3}{2}}}$.

(b) For the straight wire, we quite obviously revert to Cartesian co-ordinates. Therefore we use $\mathrm{d}\mathbf{s} = \mathbf{i}\mathrm{d}x$, $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ so that

$$\frac{\mathrm{d}\mathbf{s} \times \mathbf{r}}{r^3} = \frac{y\mathrm{d}x\mathbf{k} - z\mathrm{d}x\mathbf{j}}{(x^2 + y^2 + z^2)^{\frac{3}{2}}} \text{ and therefore}$$
$$\int_C \frac{\mathrm{d}\mathbf{s} \times \mathbf{r}}{r^3} = (y\mathbf{k} - z\mathbf{j})\int_{-\infty}^{\infty} \frac{\mathrm{d}x}{(x^2 + y^2 + z^2)^{\frac{3}{2}}}$$

In order to evaluate this integral by hand, we let $a^2 = y^2 + z^2$ for convenience (all are constant, so we can call $y^2 + z^2$ anything) and use substitution as follows:

$$\int_{-\infty}^{\infty} \frac{\mathrm{d}x}{(a^2 + x^2)^{\frac{3}{2}}} = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{a\sec^2\phi}{a^3\sec^3\phi}\,\mathrm{d}\phi \text{ writing } x = a\sec\phi$$
$$= \frac{1}{a^2}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos\phi\mathrm{d}\phi = \frac{2}{a^2} = \frac{2}{y^2 + z^2}$$

whence
$$I\int_C \frac{\mathrm{d}\mathbf{s} \times \mathbf{r}}{r^3} = \frac{2I(y\mathbf{k} - z\mathbf{j})}{y^2 + z^2}$$

Example 8.6

A coil of wire occupies the circle $x^2 + y^2 = a^2$. Calculate the magnetic induction $\mathbf{B}$ at an arbitrary point on the z-axis due to a current I flowing in the coil, if

$$\mathbf{B} = I\int_C \frac{\mathrm{d}\mathbf{s} \times \mathbf{p}}{|\mathbf{p}|^3}$$

where $\mathbf{p}$ is a vector joining an arbitrary point on the z-axis to a point on the coil and C is the circle $x^2 + y^2 = a^2$.

Solution

The last problem was practice, but this problem contains a precise statement of the Biot–Savart Law. The integral is, however, almost always impossible to evaluate except for a few special cases of which this is one. The vector $\mathbf{p}$ is shown in Figure 8.5.

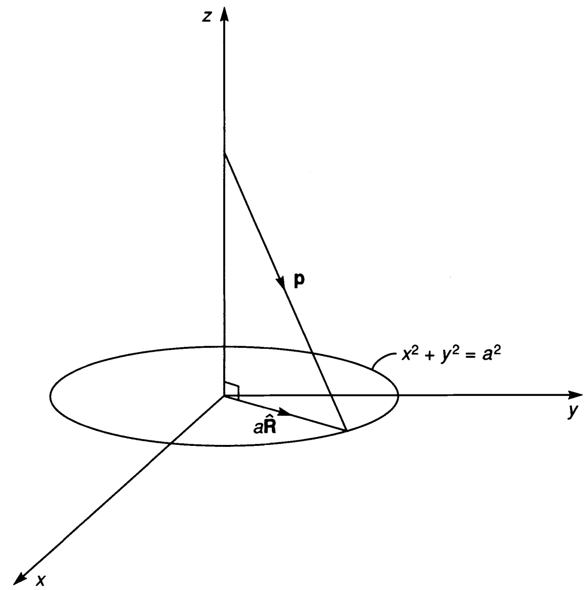


Figure 8.5 The coil of wire at $x^2 + y^2 = a^2$, and the vector $\mathbf{p}$.

The details follow those of Example 8.5, that is $\mathrm{d}\mathbf{s} = a\mathrm{d}\theta\hat{\boldsymbol{\theta}}$ and $\mathbf{p} = -z\hat{\mathbf{k}} + a\hat{\mathbf{R}}$ from the right-angled triangle drawn in Figure 8.5. Therefore

$$\frac{\mathrm{d}\mathbf{s} \times \mathbf{p}}{p^3} = \frac{-az\hat{\mathbf{R}} - a^2\hat{\mathbf{k}}}{(z^2 + a^2)^{\frac{3}{2}}}\mathrm{d}\theta$$

Now, $\qquad \hat{\mathbf{R}} = \mathbf{i}\cos\theta + \mathbf{j}\sin\theta$, hence

$$\mathbf{B} = I\int_C \frac{\mathrm{d}\mathbf{s} \times \mathbf{p}}{|\mathbf{p}|^3} = I\int_C \frac{-az\cos\theta\mathbf{i} - az\sin\theta\mathbf{j} - a^2\mathbf{k}}{(z^2 + a^2)^{\frac{3}{2}}}\mathrm{d}\theta = -\frac{2\pi a^2 I}{(z^2 + a^2)^{\frac{3}{2}}}\mathbf{k}$$

(The first two terms integrate to zero.) An important feature of this type of problem is the interpretation in terms of electromagnetism. In most real problems of this type, the integrals are very hard to evaluate and require advanced analytical techniques (such as the use of complex variables, special functions and series expansions for flat and circular geometries). Ultimately in the engineering environment, numerical evaluation is necessary.

Example 8.7 If $\mathbf{a} = (3x + y)\mathbf{i} - x\mathbf{j} + (y - 2)\mathbf{k}$, $\mathbf{b} = 2\mathbf{i} - 3\mathbf{j} + \mathbf{k}$, evaluate $\oint_C (\mathbf{a} \times \mathbf{b}) \times \mathrm{d}\mathbf{r}$ where C is the circle $x^2 + y^2 = 4$, $z = 0$.

Solution This is a routine calculation that can be made a lot easier by using the expansion of the triple vector product $(\mathbf{a} \times \mathbf{b}) \times \mathrm{d}\mathbf{r} = \mathbf{b}(\mathbf{a} \cdot \mathrm{d}\mathbf{r}) - \mathbf{a}(\mathbf{b} \cdot \mathrm{d}\mathbf{r})$. Using the expressions in the question,

$$\mathbf{a} \cdot \mathrm{d}\mathbf{r} = (3x + y)\mathrm{d}x - x\mathrm{d}y + (y - 2)\mathrm{d}z$$

and $\qquad \mathbf{b} \cdot \mathrm{d}\mathbf{r} = 2\mathrm{d}x - 3\mathrm{d}y + \mathrm{d}z$

whence

$\mathbf{b}(\mathbf{a} \cdot \mathrm{d}\mathbf{r}) - \mathbf{a}(\mathbf{b} \cdot \mathrm{d}\mathbf{r}) = \mathbf{i}(2\mathbf{a} \cdot \mathrm{d}\mathbf{r} - (3x + y)\mathbf{b} \cdot \mathrm{d}\mathbf{r}) + \mathbf{j}(-3\mathbf{a} \cdot \mathrm{d}\mathbf{r} + x\mathbf{b} \cdot \mathrm{d}\mathbf{r}) + \mathbf{k}(\mathbf{a} \cdot \mathrm{d}\mathbf{r} - (y - 2)\,\mathbf{b} \cdot \mathrm{d}\mathbf{r})$. On C, $x = 2\cos\theta$, $y = 2\sin\theta$, $z = 0$ so $\mathbf{a} \cdot \mathrm{d}\mathbf{r} = (12\sin\theta\cos\theta - 4)\mathrm{d}\theta$ and $\mathbf{b} \cdot \mathrm{d}\mathbf{r} = (-4\sin\theta - 6\cos\theta)\mathrm{d}\theta$. Substituting these expressions into $\mathbf{b}(\mathbf{a} \cdot \mathrm{d}\mathbf{r}) - \mathbf{a}(\mathbf{b} \cdot \mathrm{d}\mathbf{r})$ and tidying up gives

$$\begin{aligned}(\mathbf{a} \times \mathbf{b}) \times \mathrm{d}\mathbf{r} &= \mathbf{i}(60\sin\theta\cos\theta + 28\cos^2\theta)\mathrm{d}\theta + \mathbf{j}(-44\sin\theta - 12\cos^2\theta + 12)\mathrm{d}\theta \\ &\quad + \mathbf{k}(24\cos\theta\sin\theta + 8\sin^2\theta - 4 - 8\sin\theta - 12\cos\theta)\mathrm{d}\theta\end{aligned}$$

so

$$\oint_C (\mathbf{a} \times \mathbf{b}) \times d\mathbf{r} = \int_0^{2\pi} \mathbf{i}(60\sin\theta\cos\theta + 28\cos^2\theta) + \mathbf{j}(-44\sin\theta - 12\cos^2\theta + 12)$$
$$+ \mathbf{k}(24\cos\theta\sin\theta + 8\sin^2\theta - 4 - 8\sin\theta - 12\cos\theta)d\theta$$
$$= 28\pi\mathbf{i} + 12\pi\mathbf{j}$$

We have used the results $\int_0^{2\pi} \cos\theta\sin\theta d\theta = 0$, $\int_0^{2\pi} \cos^2\theta d\theta = \pi$ (whereas you could have used computer algebra!)

Example 8.8 A vector field of force is given by

$$\mathbf{F} = (y^2\sin x + z^3)\mathbf{i} + (2y\sin x - 4)\mathbf{j} + (3xz^2 + 2)\mathbf{k}$$

Find the work done in moving a mass m in this field from the point $(0, 1, -1)$ to $\left(\frac{\pi}{2}, -1, 2\right)$.

Solution The work done in moving a particle of mass m from $(0, 1, -1)$ to $\left(\frac{\pi}{2}, -1, 2\right)$ along the path C is given by $\int_C \mathbf{F} \cdot d\mathbf{r}$. The question does not give the path C therefore there is a hint that this vector field $\mathbf{F}$ is *conservative*, in which case the work done would be independent of the particular path taken. If $\mathbf{F}$ is conservative, then $\nabla \times \mathbf{F} = \mathbf{0}$ and there exists a scalar potential ϕ such that $\mathbf{F} = \nabla\phi$. Hence in this case

$$\int_C \mathbf{F} \cdot d\mathbf{r} = \int_C \nabla\phi \cdot d\mathbf{r} = \int_C d\phi = \phi_A - \phi_B$$

where A and B are the end points of the path C. First of all we check whether $\nabla \times \mathbf{F} = \mathbf{0}$ as follows:

$$\nabla \times \mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ y^2\cos x + z^3 & 2y\sin x - 4 & 3xz^2 + 2 \end{vmatrix}$$
$$= \mathbf{i}(0 - 0) + \mathbf{j}(3z^2 - 3z^2) + \mathbf{k}(2y\cos x - 2y\cos x)$$
$$= \mathbf{0}$$

Hence $\mathbf{F}$ is indeed conservative. We now need to find the scalar potential ϕ such that $\mathbf{F} = \nabla\phi$, that is, $\dfrac{\partial\phi}{\partial x} = (y^2\sin x + z^3)$, $\dfrac{\partial\phi}{\partial y} = (2y\sin x - 4)$, $\dfrac{\partial\phi}{\partial z} = (3xz^2 + 2)$. Integrating each expression gives:

$$\phi = y^2 \sin x + xz^3 + f_1(y, z)$$
$$\phi = y^2 \sin x - 4y + f_2(z, x)$$
$$\phi = y^2 \sin x + 2z + f_3(x, y)$$

Comparing all three of these expressions for ϕ leads to

$$\phi = y^2 \sin x + xz^3 - 4y + 2z + k \text{ where } k \text{ is a constant.}$$

At $(0, 1, 1)$, $\phi = -4 + 2 + k = k - 2$.

At $\left(\frac{\pi}{2}, -1, 2\right)$, $\phi = 1 + 4\pi + 4 + 8 + k = 4\pi + 13 + k$.

Hence $\int_C \mathbf{F} \cdot d\mathbf{r} = \phi\left(\frac{\pi}{2}, -1, 2\right) - \phi(0, 1, 1) = 4\pi + 13 + k - k + 2$
$$= 4\pi + 15.$$

Example 8.9 Show that the work done by a unit mass moving along a curve C under the action of a force $\mathbf{F}$ is $\int_C \mathbf{F} \cdot \hat{\mathbf{T}} ds$ where $\hat{\mathbf{T}}$ is the unit tangent. Hence or otherwise find the work done by a unit mass moving along the following curves C under the given forces $\mathbf{F}$:
(a) $\mathbf{F} = \hat{\mathbf{T}}\sqrt{x^2 + y^2}$, C is a semi-circle parameterised by $x = 1 - \cos\theta$, $y = \sin\theta$, $0 \le \theta \le \pi$.
(b) $\mathbf{F} = z\hat{\mathbf{k}}$, C is the curve $x = f_1(t)$, $y = f_2(t)$, $z = \sqrt{1 + t^2}$, $0 \le t \le 1$.
(c) $\mathbf{F} = x^3\mathbf{i} + \mathbf{j} + \sin ky\mathbf{k}$, C is the curve $x = 1$, $y = \sinh^{-1} t$, $z = \sqrt{1 + t^2}$, $0 \le t \le 1$.
(d) $\mathbf{F} = mg\hat{\mathbf{k}}$ along the same curve as in part (c).

Solution The work done by a unit mass moving along a curve C under the action of a force $\mathbf{F}$ is the integral $\int_C \mathbf{F} \cdot d\mathbf{r}$ which can be written $\int_{t_0}^{t_1} \mathbf{F} \cdot \frac{d\mathbf{r}}{dt} dt$ where $t_0 \le t \le t_1$ is the parameter range that defines the curve C ($\mathbf{r} = \mathbf{r}(t)$ is the parametric form of C). Also, we can further write this integral as

$$\int_C \mathbf{F} \cdot \frac{d\mathbf{r}}{dt} \cdot \frac{dt}{ds} ds$$

where s is the arc length. Moreover, the unit tangent $\hat{\mathbf{T}}$ is defined by

$$\hat{\mathbf{T}} = \frac{d\mathbf{r}}{dt}\Big/\left|\frac{d\mathbf{r}}{dt}\right| = \frac{d\mathbf{r}}{dt}\Big/\frac{ds}{dt} = \frac{d\mathbf{r}}{ds}$$

and thus $$\int_C \mathbf{F} \cdot d\mathbf{r} = \text{work done} = \int_C \mathbf{F} \cdot \hat{\mathbf{T}} ds \text{ as required.}$$

(a) With $\mathbf{F} = \hat{\mathbf{T}}\sqrt{x^2 + y^2}$, $\mathbf{F} \cdot \hat{\mathbf{T}} = \sqrt{x^2 + y^2}$. On C
$x^2 + y^2 = (1 - \cos\theta)^2 + \sin^2\theta = 2 - 2\cos\theta = 4\sin^2 \frac{1}{2}\theta$
Thus, on C, $\sqrt{x^2 + y^2} = 2\sin\frac{1}{2}\theta$.
We therefore have that

$$\int_C \mathbf{F} \cdot \hat{\mathbf{T}} ds = \int_0^\pi 2\sin\tfrac{1}{2}\theta d\theta \text{ since, on } C \; ds = |d\mathbf{r}| = 1 \cdot d\theta = d\theta$$
$$= \left[-4\cos\tfrac{1}{2}\theta\right]_0^\pi = 4$$

(b) For $\mathbf{F} = z\hat{\mathbf{k}}$, we use the definition of work done, $\int_C \mathbf{F} \cdot d\mathbf{r}$ as there is no advantage in working with the unit tangent. On C

$$d\mathbf{r} = f_1'(t)dt\mathbf{i} + f_2'(t)dt\mathbf{j} + \frac{tdt}{\sqrt{1 + t^2}}\mathbf{k}$$

so $\mathbf{F} \cdot d\mathbf{r} = \sqrt{1 + t^2} \cdot \frac{tdt}{\sqrt{1 + t^2}} = tdt$

Thus, $\int_C \mathbf{F} \cdot d\mathbf{r} = \int_0^1 tdt = \frac{1}{2}$. (Note that the functions $f_1(t)$ and $f_2(t)$ do not feature since $\mathbf{F} = z\hat{\mathbf{k}}$ is everywhere perpendicular to the x–y plane.)
(c) Again for this part of the question, there is no need to use $\hat{\mathbf{T}}$, so we proceed as in the last part. On C, $\mathbf{F} = \mathbf{i} + \mathbf{j} + t\mathbf{k}$, and $d\mathbf{r} = \frac{1}{\sqrt{1 + t^2}}\mathbf{j} + \frac{t}{\sqrt{1 + t^2}}\mathbf{k}$
so $\mathbf{F} \cdot d\mathbf{r} = \frac{1}{\sqrt{1 + t^2}} dt + \frac{t^2}{\sqrt{1 + t^2}} dt = \sqrt{1 + t^2}\, dt$ whence
$\int_C \mathbf{F} \cdot d\mathbf{r} = \int_0^1 \sqrt{1 + t^2}\, dt$. This integral is a little awkward to evaluate, so use computer algebra if you are squeamish. It yields to the substitution $t = \sinh\theta$. Here are the details:

$$\int_0^1 \sqrt{1 + t^2}\, dt = \int_0^{\sinh^{-1}(1)} \cosh^2\theta d\theta$$
$$= \frac{1}{2}\int_0^{\sinh^{-1}(1)} (1 + \cosh 2\theta) d\theta$$

$$= \frac{1}{2}\left[\theta + \tfrac{1}{2}\sinh 2\theta\right]_0^{\sinh^{-1}(1)}$$

$$= \frac{1}{2}\left[\sinh^{-1} 1 + 1\sqrt{1-1}\right] \text{ (using } \sinh 2\theta = 2\cosh\theta\sinh\theta)$$

$$= \frac{1}{2}\sinh^{-1} 1$$

(d) If $\mathbf{F} = mg\hat{\mathbf{k}}$, then this represents the force due to gravity. The curve $x = 1$, $y = \sinh^{-1} t$, $z = \sqrt{1 + t^2}$ is the *catenary* $z = \cosh y$. This problem is therefore one of calculating the work done by a bead of mass m sliding down a frictionless wire shaped like a catenary from the point $z = 1$ ($t = 0$) to $z = \sqrt{2}$ ($t = 1$). Hence

$\int_C \mathbf{F} \cdot d\mathbf{r} = \int_0^1 \frac{mgt}{\sqrt{1 + t^2}}\, dt = mg\left[\sqrt{1 + t^2}\right]_0^1 = mg(\sqrt{2} - 1)$ which is the expected answer for a conservative field of force.

8.3 Exercises

8.1. Evaluate the following scalar line integrals:

(a) $\int_C xy^2\, ds$ where C is the semi-circle $x^2 + y^2 = a^2$, $z = 0$, $y \geq 0$.

(b) $\int_C \phi dx$ where $\phi = x^2 + y^4 + z^6$ and C is the curve $x = t^3$, $y = t^2$, $z = t$ from the point (0, 0, 0) to the point (1, 1, 1).

(c) $\int_C (a^2 + b^2 - r^2)^{\frac{1}{2}}\, ds$ where C is the ellipse

$$x = a\cos\theta,\ y = b\sin\theta \qquad 0 \leq \theta \leq 2\pi.$$

8.2. Evaluate the integral $\int_C \mathbf{F} \cdot d\mathbf{r}$ for the following cases:

(a) $\mathbf{F} = x\mathbf{i}$, C is the line $z = 0$, $x = y$ from (0, 0, 0) to (2, 2, 0),

(b) $\mathbf{F} = \mathbf{r}$, C is the helix ($\sin t$, $\cos t$, t) from (0, 1, 0) to (0, 1, 2π),

(c) $\mathbf{F} = \mathbf{r}$, C is the straight line $x = 0$, $y = 1$ $\quad 0 \leq z \leq 2\pi$.

Why are the solutions to parts (b) and (c) the same?

8.3. Evaluate the integral $\int_C \mathbf{F} \cdot d\mathbf{r}$ where $\mathbf{F} = 2xy\mathbf{i} + x^2\mathbf{j} + 2z\mathbf{k}$ along the paths

(a) $x = t$, $y = t^2$, $z = t^3$,

(b) the straight line segments (0, 0, 0) to (1, 0, 0), (1, 0, 0) to (1, 1, 0) and (1, 1, 0) to (1, 1, 1).

Verify that $\mathbf{F}$ is conservative and find a potential ϕ such that $\mathbf{F} = \nabla\phi$.

8.4. If $\mathbf{F}$ is the field $x^2\mathbf{i} + y^2\mathbf{j} + z^2\mathbf{k}$ and C is the twisted cubic $x = t$, $y = \frac{t^2}{\sqrt{2}}$, $z = \frac{t^3}{3}$ between the points (0, 0, 0) and $\left(1, \frac{1}{\sqrt{2}}, \frac{1}{3}\right)$, evaluate the integrals:,

(a) $\int_C \mathbf{F}ds$, $\quad$ (b) $\int_C \mathbf{F} \cdot d\mathbf{r}$, $\quad$ (c) $\int_C \mathbf{F} \times d\mathbf{r}$.

8.5. Show that the *circulation* of a constant vector $\mathbf{A}$, $\oint \mathbf{A} \cdot d\mathbf{r}$, around any closed curve must be zero.

8.6. Evaluate $\int_C \mathbf{r}ds$ and $\int_C \mathbf{r} \times d\mathbf{r}$ where C is the section of the helix ($a\cos t$, $a\sin t$, bt) from (a, 0, 0) to (a, 0, $2\pi b$).

8.7. Calculate the work done by a unit mass subject to a force $\mathbf{F} = (x - 3y)\mathbf{i} + (y - 2x)\mathbf{j}$ in going once round the ellipse $\frac{x^2}{9} + \frac{y^2}{4} = 1$ counter-clockwise.

8.8. A force $\mathbf{F}$ acts on a particle which is moving in two dimensions along the semi-circle C: $x = 1 - \cos\theta$, $y = \sin\theta$, $0 \leq 0 \leq \pi$. Find the work done in the cases when

(a) $\mathbf{F} = \mathbf{i}\sqrt{x^2 + y^2}$

(b) $\mathbf{F} = \hat{\mathbf{T}}\sqrt{x^2 + y^2}$ ($\hat{\mathbf{T}}$ is the unit tangent to C)

(*Hint*: show that $\sqrt{x^2 + y^2} = 2\sin\frac{1}{2}\theta$ on C.)

8.9. Given $\mathbf{F} = (2x - 3y)\mathbf{i} + (y - z)\mathbf{j} + x^2\mathbf{k}$, $\mathbf{G} = y\mathbf{i} + x\mathbf{j}$ evaluate $\int_C (\mathbf{F} \times \mathbf{G}) \times d\mathbf{r}$

where C is the circle $x^2 + y^2 = a^2$, $z = 0$.

8.10. Find the value of $\int_C \hat{\mathbf{T}} \cdot d\mathbf{r}$ where $\hat{\mathbf{T}}$ is the unit tangent to the curve C in the following cases:

(a) C is the circle $x^2 + y^2 = a^2$, $z = 0$

(b) C is the square with vertices (0, 0), (0, 1), (1, 0), (1, 1)

(c) C is the triangle with vertices (0, 0), (0, 1), (1, 0).

8.11. The formula $\mathbf{B} = I\int_C \frac{d\mathbf{s} \times \mathbf{p}}{|p|^3}$ calculates the magnetic induction $\mathbf{B}$ due to a current I flowing in the coil of shape C. $\mathbf{p}$ is a vector joining an arbitrary point on the z-axis (see Example 8.6). Write down but *do not evaluate* $\mathbf{B}$ if C is the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ ($x = a\cos\theta$, $y = b\sin\theta$). Your answer should be an integral of a vector function of θ with limits 0 and 2π.

Topic Guide

Double and Triple Integrals
Change of Order
Change of Variable
Jacobian
Green's Theorem in the Plane

9 Multiple Integration

9.1 Fact Sheet

Multiple integrals considered here are of two types. *Double integrals* where two integrations are performed and *Triple integrals* where three integrations are performed. The notion is easily extended, but we concentrate here on two and three integrations as these have direct application to two- and three-dimensional problems.

A double integral in variables x and y is written

$$\int_D \int f(x, y)\mathrm{d}x\mathrm{d}y$$

where $f(x, y)$ is a continuous function of x and y throughout the domain D. D is a region of the x–y plane which is closed and simply connected. Figure 9.1 demonstrates what happens when a double integral is evaluated. We decide whether to integrate with respect to x or y first, the choice is arbitrary (but select which is easier). Sometimes it proves only possible to integrate with respect to either x or y first, in this case the choice is made for you.

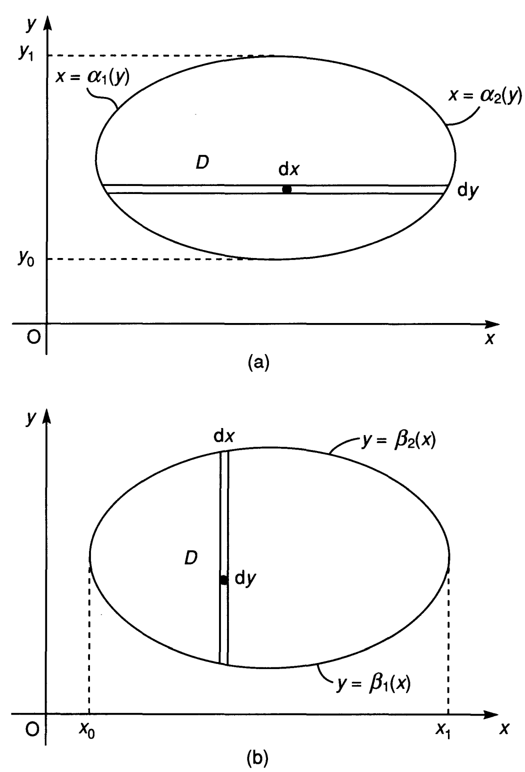


Figure 9.1 Double integration: (a) with respect to x first, (b) with respect to y first.

Integrating with respect to x first, we take the element of area dxdy, which may be thought of as a pixel on a screen, and multiply it by the local value of $f(x, y)$. The quantity $f(x, y)\mathrm{d}x\mathrm{d}y$ is then summed along the strip shown in Figure 9.1(a). As y is not changing along this strip, this amounts to integrating $f(x, y)$ with respect to x keeping y constant. The limits are $x = \alpha_1(y)$ and $x = \alpha_2(y)$, the curves that describe the border of D. Once this integration is done, all traces of x should have disappeared. What we have, in effect, is a weighted length of the strip shown in Figure 9.1(a). This weighted strip is now scanned from $y = y_0$ to $y = y_1$, that is an integration is performed with respect to y from $y = y_0$ to $y = y_1$. The result of the evaluation of

$$\int_{y_0}^{y_1} \int_{x=\alpha_1(y)}^{x=\alpha_2(y)} f(x, y)\mathrm{d}x\mathrm{d}y$$

In a similar fashion, if we integrate with respect to y first we sum $f(x, y)\mathrm{d}x\mathrm{d}y$ along vertical strips. The limits must now be written $y = \beta_1(x)$, $y = \beta_2(x)$ so all reference to y has gone once the 'weighted length' strip is calculated. These strips are then summed from x_0 to x_1 to cover D.

This double integral is written

$$\int_{x_0}^{x_1} \int_{y=\beta_1(x)}^{y=\beta_2(x)} f(x, y)\mathrm{d}y\mathrm{d}x$$

Note in particular the order dydx here. The notation adopted is to put dydx in the order of integration so that

$$\int_{x_0}^{x_1} \int_{y=\beta_1(x)}^{y=\beta_2(x)} f(x, y)\mathrm{d}y\mathrm{d}x = \int_{x_0}^{x_1} \left\{ \int_{y=\beta_1(x)}^{y=\beta_2(x)} f(x, y)\mathrm{d}y \right\} \mathrm{d}x$$

The brace is solely a function of x. Other books adopt different notations, so the reader is warned!

Often it is not convenient to use Cartesian co-ordinates. Most often, the alternative is plane polars, but if the transformation can be written

$$u=u(x, y),\ v=v(x, y)$$

then we use

$$\mathrm{d}u\mathrm{d}v = \frac{\partial(u, v)}{\partial(x, y)}\,\mathrm{d}x\mathrm{d}y$$

where $\dfrac{\partial(u, v)}{\partial(x, y)} = \begin{vmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \\ \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \end{vmatrix}$ is the usual Jacobian (see Chapter 2). The domain of the integral in the u–v space needs to be determined, then the integration can be performed in the usual way. Any problems in the way of singularities are signalled by the vanishing of the Jacobian (see Example 2.8).

Triple integration is a direct extension of double integration. An integral of $f(x, y, z)$ performed in the order z, y, x will be computed as follows: First $f(x, y, z)$ is integrated with respect to z holding both x and y fixed. The limits will (in general) depend on x and y. We then have an integral only involving x and y, and this is performed as a double integral. The notation for triple integration follows that for double integrals:

$$\int_a^b \int_{p_1(x)}^{p_2(x)} \int_{q_1(x,y)}^{q_2(x,y)} f(x, y, z)\mathrm{d}z\mathrm{d}y\mathrm{d}x = \int_a^b \left[\int_{p_1(x)}^{p_2(x)} \left\{ \int_{q_1(x,y)}^{q_2(x,y)} f(x, y, z)\mathrm{d}z \right\} \mathrm{d}y \right] \mathrm{d}x$$

A change of variable is again a straightforward extension of the two-dimensional case: if $u = u(x, y, z)$, $v = v(x, y, z)$, $w = w\,(x, y, z)$ then

$$\mathrm{d}u\mathrm{d}v\mathrm{d}w = \frac{\partial(u, v, w)}{\partial(x, y, z)} \mathrm{d}x\mathrm{d}y\mathrm{d}z$$

so that

$$\iiint_{V_2} g(u, v, w)\mathrm{d}u\mathrm{d}v\mathrm{d}w = \iiint_{V_1} f(x, y, z) \frac{\partial(u, v, w)}{\partial(x, y, z)} \mathrm{d}x\mathrm{d}y\mathrm{d}z$$

where
$$\begin{aligned} g(u, v, w) &= g(u(x, y, z), v(x, y, z), w(x, y, z)) \\ &= f(x, y, z) \end{aligned}$$

and

$$\frac{\partial(u, v, w)}{\partial(x, y, z)} = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} & \frac{\mathrm{d}v}{\partial z} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} & \frac{\partial w}{\partial z} \end{vmatrix}$$

There are many applications of multiple integrals; two of the commonest are finding centres of mass and finding moments of inertia (see Examples 9.13, 9.16 and 9.17). There are also applications to statistics in the field of marginal probability, but this is outside the scope of this Work Out.

9.2 Worked Examples

Example 9.1 Evaluate the double integral $\int_D\int(x^2 + y)\mathrm{d}x\mathrm{d}y$ where D is the rectangle in the x–y plane bounded by the lines $x = 1, 2$; $y = 0, 2$. Repeat the exercise, reversing the order of integration.

Solution The order $\mathrm{d}x\mathrm{d}y$ implies (in the convention adopted here) that x is integrated first. That is, y is held constant. Figure 9.2 shows a typical strip and the integral is $\int_0^2\int_0^1 (x^2 + y)\mathrm{d}x\mathrm{d}y$, the limits in this case being very straightforward to determine since they are all constant. Thus

$$\begin{aligned} \int_0^2\int_0^1 (x^2 + y)\mathrm{d}x\mathrm{d}y &= \int_0^2 \left[\frac{x^3}{3} + yx\right]_0^1 \mathrm{d}y \\ &= \int_0^2\left(\frac{1}{3} + y\right)\mathrm{d}y \\ &= \left[\frac{1}{3}y + \frac{1}{2}y^2\right]_0^2 = \frac{2}{3} + 2 = \frac{8}{3} \end{aligned}$$

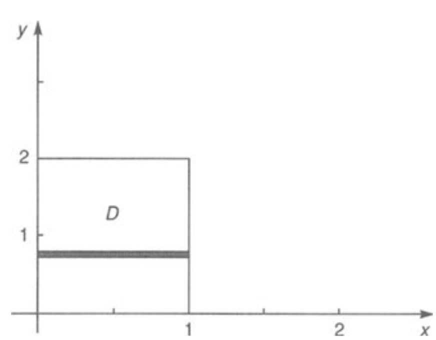


Figure 9.2 Integrating with respect to x first.

Reversing the order of integration is very straightforward when, as is the case here, the limits are constant. Figure 9.3 shows a typical strip.

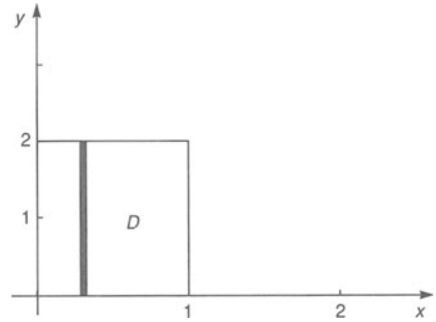


Figure 9.3 Integrating with respect to y first.

The calculation goes as follows
$$\int_0^1\int_0^2 (x^2 + y)\mathrm{d}y\mathrm{d}x = \int_0^1 [x^2y + \tfrac{1}{2}y^2]_0^2\,\mathrm{d}x$$
$$= \int_0^1 (2x^2 + 2)\mathrm{d}x$$
$$= [\tfrac{2}{3}x^3 + 2x]_0^1$$
$$= \frac{2}{3} + 2 = \frac{8}{3}$$

The same answer as before. When the limits are not constant (see Example 9.3) swapping the order of integration is more difficult and a diagram becomes virtually essential.

Example 9.2 Find the mass of a plate the shape of which is identical to the area in the positive quadrant bounded by the lines $y = 2x + 1$, $y = x^2 + 1$ and whose density ρ at the point (x, y) is given by the expression x^2y.

Solution The mass of a small element of the plate is $\rho\mathrm{d}x\mathrm{d}y$, that is $x^2y\mathrm{d}x\mathrm{d}y$. Its total mass is therefore $\int_D\int x^2y\mathrm{d}x\mathrm{d}y$ or $\int_D\int x^2y\mathrm{d}y\mathrm{d}x$ depending on the order of the integration. Figure 9.4 illustrates the area D.

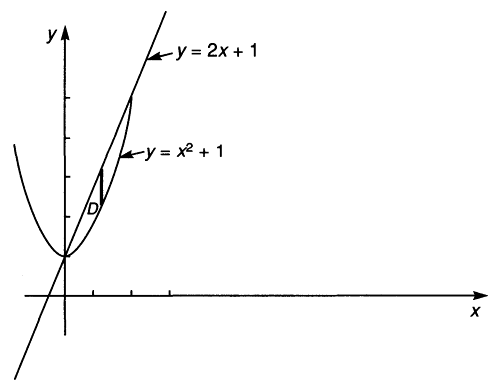


Figure 9.4 The region D between the curves $y = 2x + 1$ and $y = x^2 + 1$.

If we take vertical strips as indicated, then we evaluate $\int_D\int x^2y\mathrm{d}y\mathrm{d}x$. The limits for y are $y = x^2 + 1$ (lower) and $y = 2x + 1$ (upper). These two curves intersect where $x^2 + 1 = 2x + 1$, that is $x = 0, 2$ so that $y = 1, 5$ respectively at these points. Hence we evaluate $\int_0^2\int_{x^2+1}^{2x+1} x^2y\mathrm{d}y\mathrm{d}x = \int_0^2 \left[\tfrac{1}{2}x^2y^2\right]_{x^2+1}^{2x+1}\mathrm{d}x = \int_0^2 \tfrac{1}{2}x^2(2x + 1)^2 - \tfrac{1}{2}x^2(x^2 + 1)^2\,\mathrm{d}x$
$$= \int_0^2 -\tfrac{1}{2}x^6 + x^4 + 2x^3\mathrm{d}x$$
$$= \left[-\frac{x^7}{14} + \frac{x^5}{5} + \frac{2}{4}x^4\right]_0^2$$
$$= -\frac{128}{14} + \frac{32}{5} + 8 = \frac{184}{35}$$

Example 9.3 Reverse the order of integration for the integral in Example 9.2, $\int_0^2\int_{x^2+1}^{2x+1} x^2y\,dy\,dx$.

Solution As stated in Example 9.1, reversal of order of integration is not straightforward unless the limits are constant. They are not constant in this example. We have already drawn the area in Figure 9.4 and this is an essential prelude for evaluating double integrals, especially when reversing the order. Figure 9.5 shows the area of integration and we can deduce the limits.

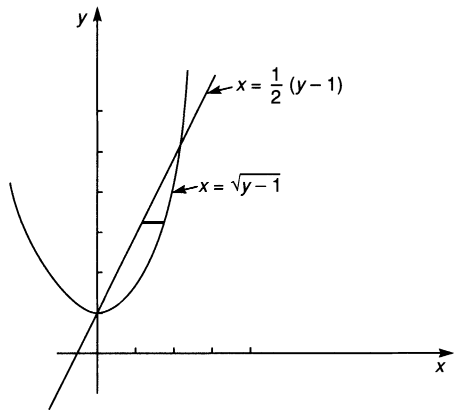


Figure 9.5 The same area as in Figure 9.4, but with 'horizontal' integration first.

The bordering lines are $y = 2x + 1$ and $y = x^2 + 1$ or $x = \frac{1}{2}(y - 1)$ and $x = \sqrt{y-1}$ (since $x > 0$, the positive square root is taken, but be mindful of square roots and check which root is appropriate to the problem). Thus the double integral is

$$\int_1^5\int_{\frac{1}{2}(y-1)}^{\sqrt{y-1}} x^2y\,dx\,dy = \int_1^5 \left[\tfrac{1}{3}x^3y\right]_{\frac{1}{2}(y-1)}^{\sqrt{y-1}} dy$$
$$= \int_1^5 \tfrac{1}{3}y(y-1)\sqrt{y-1} - \tfrac{1}{24}(y-1)^3\,y\,dy$$

These integrals are elementary to evaluate, but very tedious, so are fodder for computer algebra packages such as DERIVE©; however here are the details for those who need to see them.

$$\int_1^5 \tfrac{1}{3}y(y-1)\sqrt{y-1} - \tfrac{1}{24}(y-1)^3\,y\,dy$$
$$= \int_1^5 \tfrac{1}{3}(y-1+1)(y-1)\sqrt{y-1} - \tfrac{1}{24}(y-1)^3(y-1+1)dy$$
$$= \int_1^5 \tfrac{1}{3}(y-1)^{5/2} + \tfrac{1}{3}(y-1)^{3/2} - \tfrac{1}{24}(y-1)^4 - \tfrac{1}{24}(y-1)^3\,dy$$
$$= \int_0^4 \tfrac{1}{3}u^{5/2} + \tfrac{1}{3}u^{3/2} - \tfrac{1}{24}u^4 - \tfrac{1}{24}u^3\,du \qquad (u = y - 1)$$
$$= \left[\tfrac{1}{3}(\tfrac{2}{7}u^{7/2} + \tfrac{2}{5}u^{5/2}) - \tfrac{1}{24}(\tfrac{1}{5}u^5 + \tfrac{1}{4}u^4)\right]_0^4$$
$$= \tfrac{1}{3}[\tfrac{256}{7} + \tfrac{64}{5}] - \tfrac{1}{24}[\tfrac{1024}{5} + 64]$$
$$= \tfrac{1}{3}[\tfrac{256}{7} + \tfrac{64}{5} - \tfrac{128}{5} - 8] = \tfrac{1}{3}(\tfrac{832}{35} - \tfrac{280}{35}) = \tfrac{184}{35}$$

which is the same result as before.

Example 9.4 Let D be the region of the x–y plane bounded by the straight lines $x + y = 1$, $x = 0$, $y = 0$. By changing the variables from (x, y) to (u, v) where $u = x - y$, $v = x + y$, evaluate the double integral $\int\int_D \cos\left(\frac{x-y}{x+y}\right)dx\,dy$.

Solution Under the laws of transformation:

$$du\,dv = \frac{\partial(u, v)}{\partial(x, y)}\,dx\,dy \text{ (see Chapter 2)}$$

Now $$\frac{\partial(u, v)}{\partial(x, y)} = \begin{vmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \\ \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \end{vmatrix} = \begin{vmatrix} 1 & -1 \\ 1 & 1 \end{vmatrix} = 2$$

so $$\mathrm{d}u\mathrm{d}v = 2\mathrm{d}x\mathrm{d}y$$

Since $\dfrac{\partial(u, v)}{\partial(x, y)} \neq 0$, the transformation is not singular. (In fact it is an example of a *linear* transformation and in this particular case, a rotation through $\dfrac{\pi}{4}$ counter-clockwise and an expansion by a factor $\sqrt{2}$.)

Under the transformation given ($u = x - y$, $v = x + y$) the region D (a triangle, see Figure 9.6) becomes the different triangle, Δ.

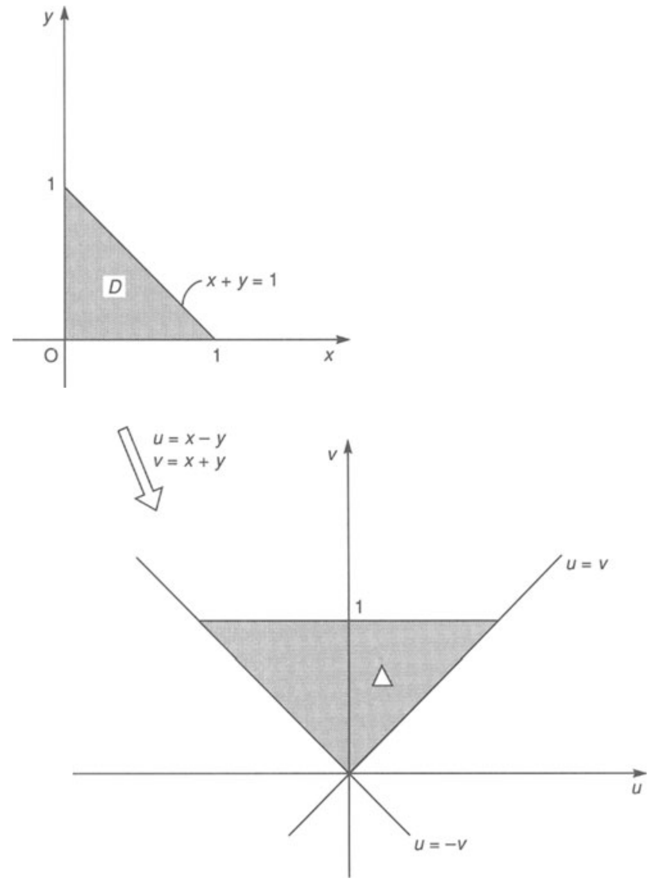


Figure 9.6 The transformed triangle.

To see this, note that by adding $u = x - y$, $v = x + y$ we get $x = \frac{1}{2}(u + v)$, and by subtracting, $y = \frac{1}{2}(v - u)$. Hence the lines $x + y = 1$, $x = 0$, $y = 0$ in the x–y plane transform into the lines $v = 1$, $u = -v$, $u = v$ respectively in the u–v plane. Also, $\cos\left(\dfrac{x - y}{x + y}\right) = \cos\left(\dfrac{u}{v}\right)$. Noting the impossibility of integrating $\cos\left(\dfrac{u}{v}\right)$ with respect to v, we choose to integrate with respect to u first. Figure 9.7 shows a typical strip and the limits.

The details of the integration are as follows:

$$\begin{aligned} \int_D\!\!\int \cos\left(\frac{x - y}{x + y}\right) \mathrm{d}x\mathrm{d}y &= \frac{1}{2}\int_\Delta\!\!\int \cos\left(\frac{u}{v}\right)\mathrm{d}u\mathrm{d}v \\ &= \frac{1}{2}\int_0^1\!\!\int_{-v}^{v} \cos\left(\frac{u}{v}\right)\mathrm{d}u\mathrm{d}v \\ &= \frac{1}{2}\int_0^1 \left[v \sin\left(\frac{u}{v}\right)\right]_{-v}^{v} \mathrm{d}v \end{aligned}$$

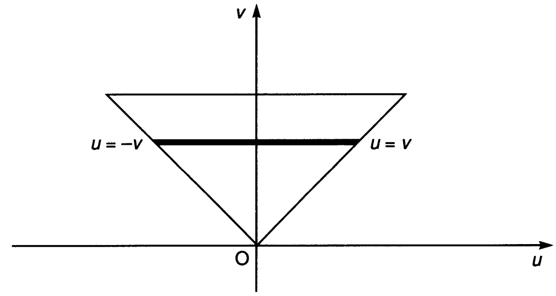


Figure 9.7 Integrating with respect to u first.

$$= \tfrac{1}{2}\int_0^1 v \sin 1 - (v \sin(-1))\mathrm{d}v$$
$$= \int_0^1 v \sin 1\mathrm{d}v \text{ since } \sin(-1) = -\sin 1$$
$$= \left[\tfrac{1}{2} v^2 \sin 1\right]_0^1$$
$$= \tfrac{1}{2} \sin 1$$

Example 9.5 Use double integration to calculate the area common to the circle $x^2 + y^2 = 4$ and the parabola $y^2 = 3x$.

Solution First of all, we need to find the common area in terms of co-ordinate geometry. The points of intersection of $x^2 + y^2 = 4$ and $y^2 = 3x$ are given by the solution of a simple quadratic and occur at the points $(1, -\sqrt{3})$ and $(1, \sqrt{3})$. (Other roots of the equation are rejected on the grounds of geometry, see Figure 9.8.)

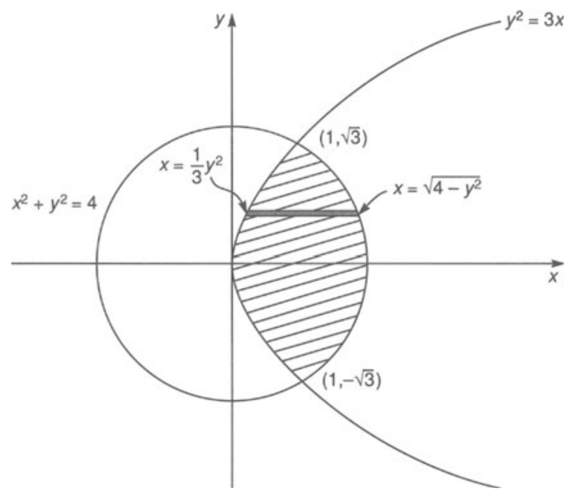


Figure 9.8 The intersection of the circle $x^2 + y^2 = 4$ and the parabola $y^2 = 3x$.

In this example, the choice of what variable to integrate with respect to first is in a sense made for us. If vertical strips were chosen, Figure 9.8 shows all too clearly that the upper and lower limits of the first integral (the ends of a vertical strip) would change part way along as they passed the intersection points $(1, -\sqrt{3})$ and $(1, \sqrt{3})$. This would entail splitting the integral into two or perhaps more parts, which is best avoided. Thus we choose horizontal strips, corresponding to integrating with respect to x first. The left-hand limit is $x = \sqrt{4 - y^2}$ and the right-hand limit is $x = \frac{1}{3}y^2$, hence the area is the double integral:

$$\int_{-\sqrt{3}}^{\sqrt{3}} \int_{\frac{1}{3}y^2}^{\sqrt{4-y^2}} \mathrm{d}x\mathrm{d}y = \int_{-\sqrt{3}}^{\sqrt{3}} (\sqrt{4 - y^2} - \tfrac{1}{3} y^2)\mathrm{d}y$$

The remaining integration involves straightforward but tedious substitution methods, unfashionable these days and often done by computer algebra. Here are the details:

$$\int_{-\sqrt{3}}^{\sqrt{3}} (\sqrt{4 - y^2} - \tfrac{1}{3} y^2)\mathrm{d}y = \int_{-\sqrt{3}}^{\sqrt{3}} (\sqrt{4 - y^2})\mathrm{d}y - \int_{-\sqrt{3}}^{\sqrt{3}} (\tfrac{1}{3} y^2)\mathrm{d}y$$

$$= \int_{-\frac{\pi}{3}}^{\frac{\pi}{3}} 4\cos^2\theta d\theta - \tfrac{1}{3}\left[\tfrac{1}{3}y^3\right]_{-\sqrt{3}}^{\sqrt{3}}$$
$$= \left[2\theta + \sin 2\theta\right]_{-\frac{\pi}{3}}^{\frac{\pi}{3}} - \tfrac{2}{9}(\sqrt{3})^3$$
$$= 2(\tfrac{2\pi}{3} + \tfrac{\sqrt{3}}{2}) - \tfrac{2}{3}\sqrt{3}$$
$$= \frac{4\pi}{3} + \frac{\sqrt{3}}{3}$$

the final answer.

Example 9.6 Evaluate the integrals $\int_D\int x^2 dxdy$, $\int_D\int y^2 dxdy$ where D is the annular domain $4 \leq x^2 + y^2 \leq 9$.

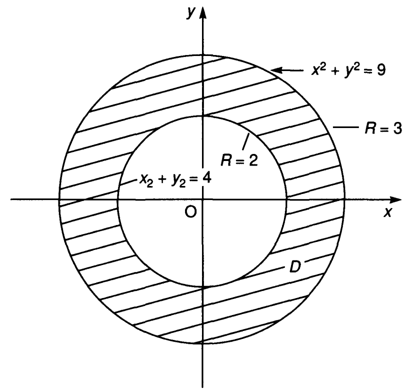


Figure 9.9 The annular region $4 \leq x^2 + y^2 \leq 9$.

Solution Figure 9.9 shows this annular region. It is not simply connected* therefore cannot be simply evaluated using Cartesian co-ordinates. Using the transformation $x = R\cos\theta$, $y = R\sin\theta$, D becomes defined by the inequality $2 \leq R \leq 3$ and also the element of area $dxdy = RdRd\theta$ so the integrals become

$$\int_0^{2\pi}\int_2^3 R^3\cos^2\theta dRd\theta \text{ and } \int_0^{2\pi}\int_2^3 R^3\sin^2\theta dRd\theta$$

Noting that $\int_2^3 R^3 dR = \left[\tfrac{1}{4}R^4\right]_2^3 = \tfrac{81}{4} - \tfrac{16}{4} = \tfrac{65}{4}$ we have:

$$\int_0^{2\pi}\int_2^3 R^3\cos^2\theta dRd\theta = \tfrac{65}{4}\int_0^{2\pi}\cos^2\theta d\theta = \frac{65\pi}{4}$$

and similarly
$$\int_0^{2\pi}\int_2^3 R^3\sin^2\theta dRd\theta = \tfrac{65}{4}\int_0^{2\pi}\sin^2\theta d\theta = \frac{65\pi}{4}$$

Example 9.7 Evaluate the double integral $\int_D\int (x^2 + y^2)dxdy$ where D is the area between the circles $x^2 + y^2 = 4$ and $x^2 + y^2 + 2x = 0$.

Solution Converting to polar co-ordinates, $x = R\cos\theta$, $y = R\sin\theta$, the circle $x^2 + y^2 = 4$ is $R = 2$ and the circle $x^2 + y^2 + 2x = 0$ is $R = -2\cos\theta$. Figure 9.10 displays the area.

Note that the use of polar co-ordinates here is more or less mandatory since it is impossible to cover the shaded area in any simple way using straight strips. It is another example of a region that is not simply connected. Note also that $R = -2\cos\theta$ implies that $\frac{\pi}{2} \leq \theta \leq \frac{3\pi}{2}$ in order to preserve $R \geq 0$. In polar co-ordinates, $x^2 + y^2 = R^2$ and $dxdy = RdRd\theta$ so that $\int_D\int (x^2 + y^2)dxdy = \int_D\int R^3 \, dRd\theta$. A little thought is required before evaluating this double integral, because, as we have already said, the formula $R = -2\cos\theta$ is only valid for $\frac{\pi}{2} \leq \theta \leq \frac{3\pi}{2}$. Outside this range for θ,

* A *simply connected* domain is one for which every closed curve that lies wholly within the domain contains only points that are in the domain itself. Put simply, it is a domain without holes.

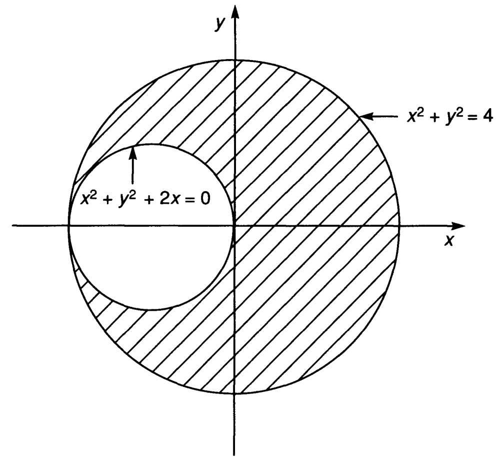


Figure 9.10 The area outside $x^2 + y^2 + 2x = 0$ but inside the circle $x^2 + y^2 = 4$.

R will vary all the way from 0 to 2. Figure 9.11 shows in diagrams the areas covered by each of the three areas in the following breakdown of the given double integral. Hence

$$\iint_D R^3 \mathrm{d}R\mathrm{d}\theta = \underset{\text{(a)}}{\int_0^{\frac{\pi}{2}}\int_0^2 R^3 \mathrm{d}R\mathrm{d}\theta} + \underset{\text{(b)}}{\int_{\frac{\pi}{2}}^{\frac{3\pi}{2}}\int_{-2\cos\theta}^2 R^3 \mathrm{d}R\mathrm{d}\theta} + \underset{\text{(c)}}{\int_{\frac{3\pi}{2}}^{2\pi}\int_0^2 R^3 \mathrm{d}R\mathrm{d}\theta}$$

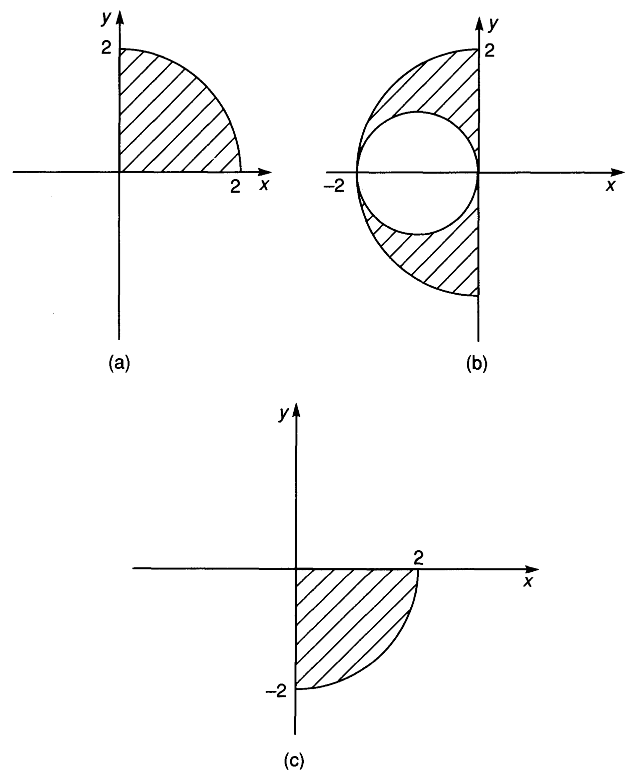


Figure 9.11 The three regions of integration corresponding to integrals (a), (b) and (c).

Evaluating each integral,

(a) $$\int_0^{\frac{\pi}{2}}\int_0^2 R^3 \mathrm{d}R\mathrm{d}\theta = \int_0^{\frac{\pi}{2}} \left[\tfrac{1}{4}R^4\right]_0^2 \mathrm{d}\theta = \int_0^{\frac{\pi}{2}} 4\mathrm{d}\theta = 2\pi$$

(b) $$\int_{\frac{\pi}{2}}^{\frac{3\pi}{2}}\int_{-2\cos\theta}^2 R^3 \mathrm{d}R\mathrm{d}\theta = \int_{\frac{\pi}{2}}^{\frac{3\pi}{2}} \left[\tfrac{1}{4}R^4\right]_{-2\cos\theta}^2 \mathrm{d}\theta = \int_{\frac{\pi}{2}}^{\frac{3\pi}{2}} (4 - 4\cos^4\theta)\mathrm{d}\theta$$
$$= \int_{\frac{\pi}{2}}^{\frac{3\pi}{2}} (\tfrac{5}{2} - 2\cos 2\theta - \tfrac{1}{2}\cos 4\theta)\mathrm{d}\theta = \tfrac{5}{2}\pi$$

(c) $$\int_{\frac{3\pi}{2}}^{2\pi}\int_0^2 R^3 \mathrm{d}R\mathrm{d}\theta = 2\pi \text{ as in part (a)}$$

Hence, the double integral is the sum of all these integrals giving

$$\int_D\int R^3 \mathrm{d}R\mathrm{d}\theta = \int_0^{\frac{\pi}{2}}\int_0^2 R^3\mathrm{d}R\mathrm{d}\theta + \int_{\frac{\pi}{2}}^{\frac{3\pi}{2}}\int_{-2\cos\theta}^{2} R^3\mathrm{d}R\mathrm{d}\theta + \int_{\frac{3\pi}{2}}^{2\pi}\int_0^2 R^3\,\mathrm{d}R\mathrm{d}\theta$$
$$= 2\pi + \tfrac{5}{2}\pi + 2\pi = \frac{13}{2}\pi$$

Example 9.8 Suppose that C is a closed contour in the x–y plane which completely encloses a domain D. Assume further that the contour C is *convex*, that is any straight line that cuts the domain, intersects C in at most two places. Let $P(x, y)$ and $Q(x, y)$ be differentiable functions in the two variables x and y throughout, and on the borders of the domain D. Then *Green's Theorem in the Plane* states that

$$\int_C P\mathrm{d}x + Q\mathrm{d}y = \int_D\int \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathrm{d}x\mathrm{d}y$$

Prove this theorem.

Solution First of all let us draw Figure 9.12 which shows both C and D and splits the contour C into left and right halves (labelled $g_1(y)$ and $g_2(y)$ respectively) and into top and bottom halves (labelled $f_2(x)$ and $f_1(x)$ respectively), and also shows two strips, horizontal and vertical.

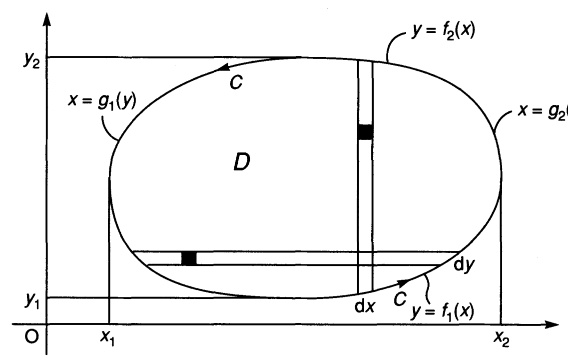


Figure 9.12 The domain D with surrounding contour C showing both vertical and horizontal strips.

Consider the integral

$$\int_D\int \frac{\partial P}{\partial y}\,\mathrm{d}y\mathrm{d}x$$

where we have selected Cartesian co-ordinates to integrate $\dfrac{\partial P}{\partial y}$, choosing to integrate y first. Using vertical strips, we apply elementary integration to give

$$\int_{x_0}^{x_1}\int_{f_1(x)}^{f_2(x)} \frac{\partial P}{\partial y}\,\mathrm{d}y\mathrm{d}x = \int_{x_0}^{x_1}\Big[P(x, y)\Big]_{y=f_1(x)}^{y=f_2(x)}\mathrm{d}x = \int_{x_0}^{x_1}(P(x, f_2) - P(x, f_1))\mathrm{d}x$$

where x_0 and x_1 denote the left and right extremities of the domain D. Now, the integrand of the last expression, $P(x, f_2) - P(x, f_1)$, is the function P evaluated on the boundary of D, that is C. Moreover, the integration with respect to x is from x_0 to x_1 so we may write $\int_{x_0}^{x_1}(P(x, f_2) - P(x, f_1))\mathrm{d}x = \int_{x_0}^{x_1}P(x, f_2)\mathrm{d}x - \int_{x_0}^{x_1}P(x, f_1)\mathrm{d}x = -\int_C P\mathrm{d}x$. The negative sign arises because C is being traversed clockwise in the middle integrals whereas the convention is to traverse integrals over a contour C *counter*-clockwise. We have thus shown that

$$\int_D\int \frac{\partial P}{\partial y}\,\mathrm{d}y\mathrm{d}x = -\int_C P\mathrm{d}x$$

In an exactly similar fashion, considering $\int\int_D \frac{\partial Q}{\partial x}\,\mathrm{d}x\mathrm{d}y$ with horizontal strips, we can show that

$$\int\int_D \frac{\partial Q}{\partial x}\,\mathrm{d}x\mathrm{d}y = \int_C Q\mathrm{d}y$$

Subtracting these two results gives us Green's Theorem in the Plane:

$$\int_C P\mathrm{d}x + Q\mathrm{d}y = \int\int_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathrm{d}x\mathrm{d}y$$

as required. There are several comments worth making here. First of all the left-hand side of the equation in this theorem upsets purists who prefer the expression

$$\int_C \left(P\frac{\mathrm{d}x}{\mathrm{d}s} + Q\frac{\mathrm{d}y}{\mathrm{d}s}\right)\mathrm{d}s = \int\int_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathrm{d}x\mathrm{d}y$$

since the left-hand side must be evaluated on C for which s is the arc length, the natural parameter. Secondly, although we have only proved the theorem for convex domains, it does hold far more generally. This is shown by dividing up the non-convex domain into suitable sub-domains. It is messy but straightforward. Thirdly, this theorem is the two-dimensional version of Stokes' Theorem (see Chapter 11), and also helps in the understanding of Cauchy's Theorem which is fundamental in the development of Complex Analysis.

Example 9.9 Use Green's Theorem in the Plane to evaluate the integral

$$\oint_C (3x^2 + y)\mathrm{d}x + (2x + y^3)\mathrm{d}y$$

where C is the circle $x^2 + y^2 = a^2$.

Solution In order to use Green's Theorem, we need to identify the functions $P(x, y)$ and $Q(x, y)$. These are $P(x, y) = 3x^2 + y$ and $Q(x, y) = 2x + y^3$, so that $\frac{\partial P}{\partial y} = 1$ and $\frac{\partial Q}{\partial x} = 2$. Green's Theorem in the Plane implies that

$$\oint_C (3x^2 + y)\mathrm{d}x + (2x + y^3)\mathrm{d}y = \int\int_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathrm{d}x\mathrm{d}y = \int\int_D \mathrm{d}x\mathrm{d}y$$

However, the right-hand side is simply the area of the domain D, that is the area of the circle $= \pi a^2$. Hence we deduce that

$$\oint_C (3x^2 + y)\mathrm{d}x + (2x + y^3)\mathrm{d}y = \pi a^2$$

The methods of Chapter 8 whereby C is parameterised by $x = a\cos\theta$, $y = a\sin\theta$ $0 \le \theta \le 2\pi$ could be used, but the algebra is much more difficult!

Example 9.10 [This example is harder than others, but is good for the use of polar co-ordinates.] Evaluate the integral $\int\int_D \frac{(x^2 + y^2)^2}{x^2y^2}\,\mathrm{d}x\mathrm{d}y$ where D is the region between the two circles $x^2 + y^2 = ax$ and $x^2 + y^2 = by$.

Solution First of all, we convert to plane polar co-ordinates $x = R\cos\theta$, $y = R\sin\theta$ and $\mathrm{d}x\mathrm{d}y = R\mathrm{d}R\mathrm{d}\theta$ (as before, see the last example). Thus the integrand becomes $\frac{(x^2 + y^2)^2}{x^2y^2} = \frac{R^4}{R^4\cos^2\theta\sin^2\theta} = 4\mathrm{cosec}^2\, 2\theta$. The region D is shown shaded in Figure 9.13.

In this problem, the first of several complications is expressing the limits of the integral succinctly. In plane polar co-ordinates, $x^2 + y^2 = ax$ is $R^2 = aR\cos\theta$ or $R = a\cos\theta$, and $x^2 + y^2$

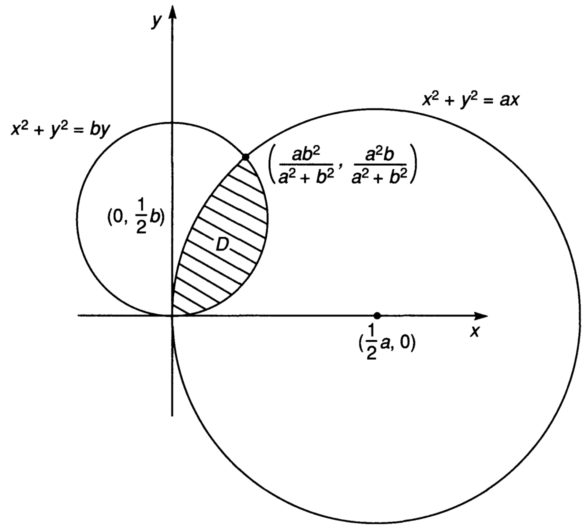


Figure 9.13 The region of intersection D.

$= by$ is $R^2 = bR\sin\theta$ or $R = b\sin\theta$. We choose to integrate with respect to θ first; the lower limit is $\theta = \sin^{-1}(R/b)$ and the upper limit is $\theta = \cos^{-1}(R/a)$. The limits on R are $R = 0$ and the intersection of $R = a\cos\theta$ and $R = b\sin\theta$. The most straightforward way of determining this intersection is to eliminate θ as follows:

$$\cos\theta = \frac{R}{a},\ \sin\theta = \frac{R}{b} \text{ are both valid at this intersection}$$

therefore $\cos^2\theta + \sin^2\theta = 1 = \dfrac{R^2}{a^2} + \dfrac{R^2}{b^2}$

whence $\qquad R^2 = \dfrac{a^2b^2}{a^2 + b^2}$

The double integral written explicitly in polar co-ordinates is thus

$$\int_0^{\frac{ab}{\sqrt{a^2+b^2}}} \int_{\sin^{-1}(R/b)}^{\cos^{-1}(R/a)} 4\mathrm{cosec}^2 2\theta\, R\mathrm{d}\theta\mathrm{d}R$$

Now, $\int 4\mathrm{cosec}^2 2\theta \mathrm{d}\theta = -2\cot 2\theta + c$ is a (slight modification of a) standard integral. Using the fact that $\cot 2\theta = \dfrac{\cos 2\theta}{\sin 2\theta} = \dfrac{\cos^2\theta - \sin^2\theta}{2\sin\theta\cos\theta} = \dfrac{1 - \tan^2\theta}{2\tan\theta}$ we see that

$$\int_{\sin^{-1}(R/b)}^{\cos^{-1}(R/a)} 4\mathrm{cosec}^2 2\theta\, R\mathrm{d}\theta = \left[\frac{\tan^2\theta - 1}{\tan\theta}\right]_{\sin^{-1}(R/b)}^{\cos^{-1}(R/a)}$$

By drawing elementary triangles, if $\cos\theta = \dfrac{R}{a}$ then $\tan\theta = \sqrt{\dfrac{a^2}{R^2} - 1}$ and if $\sin\theta = \dfrac{R}{b}$ then $\tan\theta = \dfrac{R}{\sqrt{b^2 - R^2}}$. Thus substituting the upper limit $\theta = \cos^{-1}(R/a)$ gives $\dfrac{\tan^2\theta - 1}{\tan\theta} = \dfrac{a^2 - 2R^2}{R\sqrt{a^2 - R^2}}$ and substituting the lower limit $\theta = \sin^{-1}(R/b)$ gives $\dfrac{\tan^2\theta - 1}{\tan\theta} = \dfrac{2R^2 - b^2}{R\sqrt{b^2 - R^2}}$. We thus need to evaluate

$$\int_0^{\frac{ab}{\sqrt{a^2+b^2}}} \left(\frac{a^2 - 2R^2}{R\sqrt{a^2 - R^2}} - \frac{2R^2 - b^2}{R\sqrt{b^2 - R^2}}\right) R\mathrm{d}R$$

which is evaluated term by term after cancellation of R. Let $R = a\sin\phi$ so that

$$\frac{a^2 - 2R^2}{R\sqrt{a^2 - R^2}} = \frac{a^2\cos 2\phi}{.a\cos\phi} \text{ and } \mathrm{d}R = a\cos\phi\mathrm{d}\phi$$

thus $\int \frac{a^2 - 2R^2}{\sqrt{a^2 - R^2}} \, dR = \int a^2\cos 2\phi \, d\phi = \frac{1}{2}a^2 \sin 2\phi = aR\sqrt{1 - \frac{R^2}{a^2}} = R\sqrt{a^2 - R^2}$ ignoring the arbitrary constant. The second integral is similarly $-R\sqrt{b^2 - R^2}$.

$$\text{Hence } \int_0^{\frac{ab}{\sqrt{a^2+b^2}}} \left(\frac{a^2 - 2R^2}{R\sqrt{a^2 - R^2}} - \frac{2R^2 - b^2}{R\sqrt{b^2 - R^2}}\right) R\mathrm{d}R$$

$$= \left[R\sqrt{a^2 - R^2} + R\sqrt{b^2 - R^2}\right]_0^{\frac{ab}{\sqrt{a^2+b^2}}}$$

$$= \frac{ab}{\sqrt{a^2 + b^2}}\left(\sqrt{a^2 - \frac{a^2b^2}{a^2 + b^2}} + \sqrt{b^2 - \frac{a^2b^2}{a^2 + b^2}}\right)$$

$$= \frac{ab}{\sqrt{a^2 + b^2}}\left(\frac{a^2}{\sqrt{a^2 + b^2}} + \frac{b^2}{\sqrt{a^2 + b^2}}\right)$$

$$= ab$$

Thus
$$\iint_D \frac{(x^2 + y^2)^2}{x^2 y^2} \, \mathrm{d}x\mathrm{d}y = ab$$

There are several useful tricks displayed in the midst of this solution, mainly of use to the aspiring applied mathematician. For others, there are plenty of good tests for computer algebra packages, some of which are not good with multi-valued functions such as inverse trigonometric functions.

Example 9.11 Use a suitable double integral to evaluate the improper integral $\int_0^\infty e^{-x^2} dx$.

Solution Consider the double integral $\iint_S e^{-(x^2 + y^2)} \, dS$ where S is the quarter disc $x \geq 0$, $y \geq 0$, $x^2 + y^2 \leq a^2$. Converting to polar co-ordinates (R, θ) this integral becomes $\int_0^a \int_0^{\frac{\pi}{2}} e^{-R^2} R\mathrm{d}\theta\mathrm{d}R$ where $R^2 = x^2 + y^2$, $R\cos\theta = x$ and $R\sin\theta = y$ so that $\mathrm{d}S = R\mathrm{d}\theta\mathrm{d}R$. Evaluating this integral we obtain

$$I = \frac{\pi}{2}\int Re^{-R^2} \, \mathrm{d}R = \frac{\pi}{2}\left[-\frac{1}{2}e^{-R^2}\right]_0^a = \frac{\pi}{4}(1 - e^{-a^2})$$

As $a \to \infty$, $I \to \frac{\pi}{4}$.

We now consider the double integral $I_k = \int_0^k \int_0^k e^{-(x^2+y^2)} \, \mathrm{d}x\mathrm{d}y$. The domain of this integral is a square of side k. Now, $I_k = \left(\int_0^k e^{-x^2} \, \mathrm{d}x\right)\left(\int_0^k e^{-x^2} \, \mathrm{d}x\right) = \left(\int_0^k e^{-x^2} \, \mathrm{d}x\right)^2$. A glance at Figure 9.14 will show that $I_{a/\sqrt{2}} < I < I_a$.

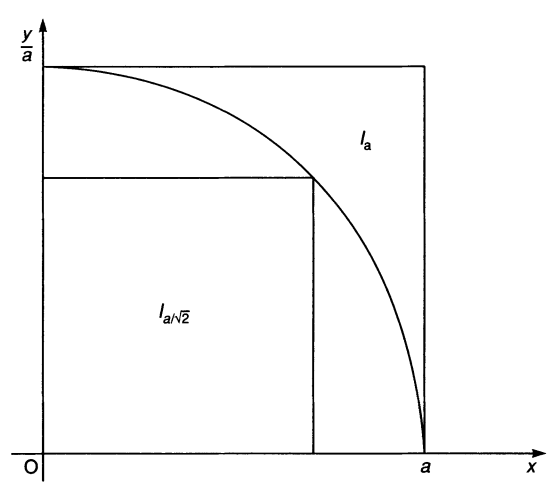


Figure 9.14 The squares over which the integrals $I_{a/\sqrt{2}}$ and I_a are taken.

However $I_k \to \left(\int_0^\infty e^{-x^2}\,dx\right)^2$ as $k \to \infty$ for all k. Hence it must be the case that $I \to \left(\int_0^\infty e^{-x^2}\,dx\right)^2$ as $a \to \infty$. Therefore $\left(\int_0^\infty e^{-x^2}\,dx\right)^2 = \frac{\pi}{4}$ and hence $\int_0^\infty e^{-x^2}\,dx = \frac{1}{2}\sqrt{\pi}$. This integral is central to statistics, being one half of the area under the bell-shaped curve usually associated with the normal distribution.

Example 9.12 Evaluate the triple integrals:

(a) $\int_0^1\int_1^2\int_2^3 xy^2z^3\,dxdydz$, (b) $\int_0^1\int_0^x\int_{x-z}^{x+z} y\,dydzdx$.

Solution There is no need to draw diagrams here since the limits are given and it is an exercise in integration only. For part (a) the limits are constant therefore the volume is a cuboid; for part (b) the limits are planes (linear expressions) thus the volume is wedge shaped. Hence for part (a):

$$\begin{aligned}
\int_0^1\int_1^2\int_2^3 xy^2z^3\,dxdydz &= \int_0^1\int_1^2 \left[\tfrac{1}{2}\,x^2y^2z^3\right]_2^3 dydz \\
&= \int_0^1\int_1^2 [\tfrac{9}{2}\,y^2z^3 - 2y^2z^3]dydz \\
&= \int_0^1 \left[\tfrac{5}{2}\,\tfrac{1}{3}\,y^3z^3\right]_1^2 dz \\
&= \tfrac{5}{2}\,\tfrac{1}{3}\int_0^1 [8 - 1]z^3dz \\
&= \tfrac{5}{6}\,\tfrac{7}{3}\left[\frac{z^4}{4}\right]_0^1 = \frac{35}{72}
\end{aligned}$$

For part (b):

$$\begin{aligned}
\int_0^1\int_0^x\int_{x-z}^{x+z} y\,dydzdx &= \int_0^1\int_0^x [\tfrac{1}{2}((x + z)^2 - (x - z)^2)]dzdx \\
&= \int_0^1\int_0^x 2xz\,dzdx \\
&= \int_0^1 \left[xz^2\right]_0^x dx \\
&= \int_0^1 x^3dx = \frac{1}{4}
\end{aligned}$$

Example 9.13 Evaluate the following three volume integrals and interpret them in terms of a real volume.

(a) $4\int_0^4\int_0^{\sqrt{16-x^2}}\int_{\frac{1}{4}(x^2+y^2)}^4 dzdydx$, (b) $4\int_0^4\int_0^{2\sqrt{z}}\int_0^{\sqrt{4z-x^2}} dydxdz$,

(c) $4\int_0^4\int_{\frac{1}{4}y^2}^4\int_0^{\sqrt{4z-y^2}} dxdzdy$.

Solution Since the limits are given we first evaluate the integrals mechanically.

$$\begin{aligned}
\text{(a) } 4\int_0^4\int_0^{\sqrt{16-x^2}}\int_{\frac{1}{4}(x^2+y^2)}^4 dzdydx &= \int_0^4\int_0^{\sqrt{16-x^2}} (4 - \tfrac{1}{4}(x^2 + y^2))dydx \\
&= \int_0^4 \left[16y - x^2y - \tfrac{1}{3}\,y^3\right]_0^{\sqrt{16-x^2}} dx \\
&= \int_0^4 [16\sqrt{16 - x^2} - x^2\sqrt{16 - x^2} - \tfrac{1}{3}(16 - x^2)^{3/2}\,dx \\
&= \int_0^4 \tfrac{2}{3}(16 - x^2)^{3/2}\,dx \\
&= \frac{512}{3}\int_0^{\frac{\pi}{2}} \cos^4\theta d\theta \quad \text{(on writing } x = 4\sin\theta) \\
&= \frac{512}{3}\int_0^{\frac{\pi}{2}} \tfrac{1}{4}(1 + \cos 2\theta)^2 d\theta \\
&= \frac{128}{3}\int_0^{\frac{\pi}{2}} \tfrac{3}{2} + 2\cos 2\theta + \tfrac{1}{2}\cos 4\theta d\theta
\end{aligned}$$

$$= \frac{128}{3}\frac{3}{2}\frac{\pi}{2} = 32\pi$$

(b) $$4\int_0^4\int_0^{2\sqrt{z}}\int_0^{\sqrt{4z-x^2}} \mathrm{d}y\mathrm{d}x\mathrm{d}z = 4\int_0^4\int_0^{2\sqrt{z}} \sqrt{4z - x^2}\,\mathrm{d}x\mathrm{d}z$$

$$= 4\int_0^4\int_0^{\frac{\pi}{2}} 4z\cos^2\theta \mathrm{d}\theta\mathrm{d}z \quad (x = 2\sqrt{z}\sin\theta)$$

$$= 16\int_0^4 \frac{\pi}{4} z\mathrm{d}z = 32\pi \quad \text{(omitting some details)}$$

(c) $$4\int_0^4\int_{\frac{1}{4}y^2}^4\int_0^{\sqrt{4z-y^2}} \mathrm{d}x\mathrm{d}z\mathrm{d}y = 4\int_0^4\int_{\frac{1}{4}y^2}^4 \sqrt{4z - y^2}\,\mathrm{d}z\mathrm{d}y$$

$$= 4\int_0^4 \left[\tfrac{1}{6}(4z - y^2)^{\frac{3}{2}}\right]_{\frac{1}{4}y^2}^4 \mathrm{d}y \quad (u^2 = 4z - y^2)$$

$$= \frac{2}{3}\int_0^4 (16 - y^2)^{3/2}\,\mathrm{d}y$$

$$= 32\pi$$

These integrals all have the same value of 32π. We are now in a position to interpret the triple integral physically. Examining the triple integral (a), in particular the limits, we see that $z = 4$, $4z = x^2 + y^2$; $x^2 + y^2 = 16$, $y = 0$; $x = 4$, $x = 0$. This leads to the volume shown in Figure 9.15.

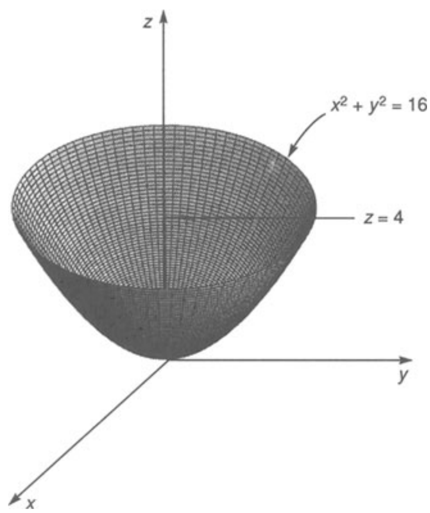


Figure 9.15 The volume represented by all three triple integrals.

You can easily check that the integrals (b) and (c) describe the same volume.

Example 9.14 A solid is formed from the intersection of the co-ordinate planes $x = 0$, $y = 0$, $z = 0$ and the plane $\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1$. Find the volume and the position of the centre of mass of this tetrahedron.

Solution Figure 9.16 shows the solid and some dimensions (it is straightforward to deduce these dimensions from the intersections of the plane $\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1$ with the axes, where x, y, z are zero in pairs). In order to find the volume, we express it as a triple integral:

$$\text{volume} = \int_0^a\int_0^{b-\frac{b}{a}x}\int_0^{c-\frac{c}{b}y-\frac{c}{a}x} \mathrm{d}z\mathrm{d}y\mathrm{d}x$$

$$= c\int_0^a\int_0^{b-\frac{b}{a}x}\left(1 - \frac{y}{b} - \frac{x}{a}\right)\mathrm{d}y\mathrm{d}x$$

$$= c\int_0^a \left[y - \frac{y^2}{2b} - \frac{xy}{a}\right]_0^{b-\frac{b}{a}x} \mathrm{d}x$$

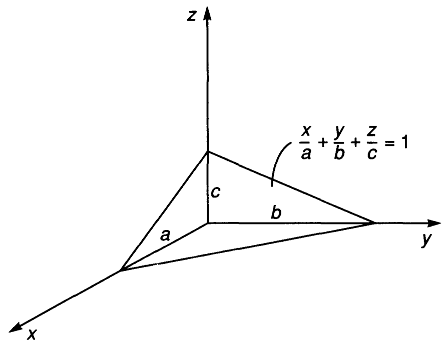


Figure 9.16 The intersection of the plane $\dfrac{x}{a} + \dfrac{y}{b} + \dfrac{z}{c} = 1$ with the three co-ordinate axes.

$$= c\int_0^a b - \frac{b}{a}x - \frac{1}{2b}\left(b - \frac{b}{a}x\right)^2 - \frac{b}{a}x\left(1 - \frac{1}{a}x\right)dx$$

$$= bc\int_0^a \left(1 - \frac{1}{a}x\right) - \frac{1}{2}\left(1 - \frac{1}{a}x\right)^2 - \frac{x}{a}\left(1 - \frac{1}{a}x\right)dx$$

so that the volume is $bc\int_0^a \frac{1}{2}\left(1 - \frac{1}{a}x\right)^2 dx$ on factorising. The integration is completed by the substitution $u = 1 - \frac{1}{a}x$ (or computer algebra of course), thus

$$\text{volume} = abc\int_0^1 \tfrac{1}{2}u^2\,du = \frac{abc}{6}$$

In order to find the centre of mass (or centroid) we use the following formula:

$$\bar{\mathbf{r}} = (\bar{x}, \bar{y}, \bar{z}) = \frac{\int_V \mathbf{r}dV}{\int_V dV}$$

where $\bar{\mathbf{r}} = (\bar{x}, \bar{y}, \bar{z})$ is the position vector of the centre of mass. Thus

$$\bar{x} = \frac{\int_V xdV}{\int_V dV} = \frac{\int_0^a\int_0^{b-\frac{b}{a}x}\int_0^{c-\frac{c}{b}y-\frac{c}{a}x} xdzdydx}{\frac{1}{6}abc}$$

where of course the limits are the same as before. Evaluating the numerator proceeds as follows:

$$\int_0^a\int_0^{b-\frac{b}{a}x}\int_0^{c-\frac{c}{b}y-\frac{c}{a}x} xdzdydx = c\int_0^a\int_0^{b-\frac{b}{a}x} x\left(1 - \frac{y}{b} - \frac{x}{a}\right)dydx$$

$$= c\int_0^a x\left[y - \frac{y^2}{2b} - \frac{xy}{a}\right]_0^{b-\frac{b}{a}x} dx$$

$$= bc\int_0^a \frac{x}{2}\left(1 - \frac{x}{a}\right)^2 dx$$

since the bracket is identical to that produced in calculating the volume.

$$\text{So, } \int_0^a\int_0^{b-\frac{b}{a}x}\int_0^{c-\frac{c}{b}y-\frac{c}{a}x} xdzdydx = bc\int_0^a \frac{x}{2} - \frac{x^2}{a} + \frac{x^3}{2a^2}dx$$

$$= bc\left[\frac{x^2}{4} - \frac{x^3}{3a} + \frac{x^4}{8a^2}\right]_0^a$$

$$= bc\left(\frac{a^2}{4} - \frac{a^2}{3} + \frac{a^2}{8}\right) = \frac{a^2bc}{24}$$

Hence,
$$\bar{x} = \frac{a^2bc/24}{abc/6} = \frac{a}{4}$$

By symmetry, $\bar{y} = \frac{b}{4}$, $\bar{z} = \frac{c}{4}$, hence the centre of mass of the tetrahedron is at the point with position vector $\left(\frac{a}{4}, \frac{b}{4}, \frac{c}{4}\right)$.

Example 9.15 Evaluate the integral $\int_V \sqrt{x^2 + y^2 + z^2}\, dV$ using cylindrical polar co-ordinates where V is the cone bounded by the curved surface $z = \sqrt{x^2 + y^2}$ and the plane $z = 3$.

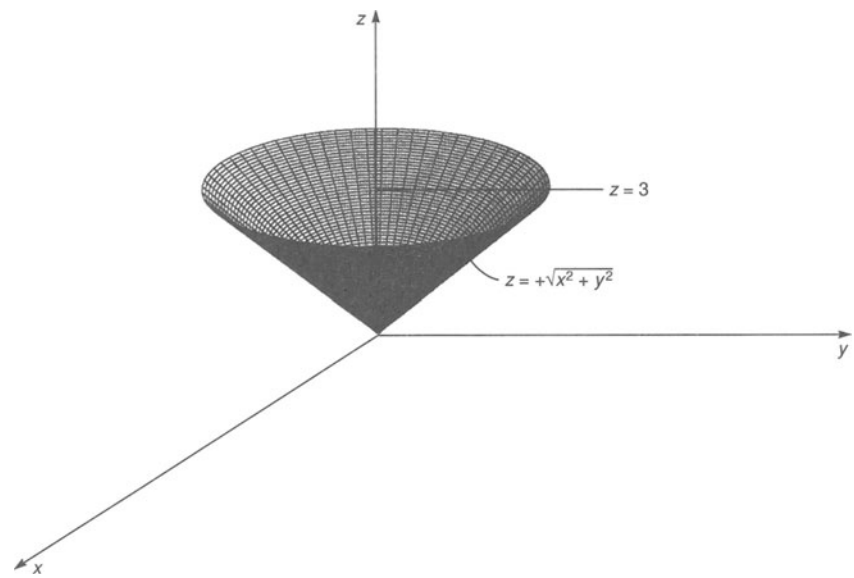


Figure 9.17 The cone $z = +\sqrt{x^2 + y^2}$ and the cut-off plane $z = 3$.

Solution The cone $z = \sqrt{x^2 + y^2}$ together with its cut-off plane $z = 3$ is shown in Figure 9.17. In cylindrical polar co-ordinates, $dV = R dR d\theta dz$. The cone is now defined by the surfaces in polar form, that is $z = R$, $z = 3$. Whence the volume integral is

$$\int_0^3 \int_0^{2\pi} \int_0^z \sqrt{z^2 + R^2}\, R dR d\theta dz$$

Note that the upper limit for R is the curved surface $R = z$. Note also that

$$\int R\sqrt{z^2 + R^2}\, dR = \tfrac{1}{3}(z^2 + R^2)^{3/2} + c$$

(Use computer algebra, or substitute $u = z^2 + R^2$ and proceed by hand.) The remainder of the evaluation is straightforward and is now given:

$$\int_0^3 \int_0^{2\pi} \int_0^z \sqrt{z^2 + R^2}\, R dR d\theta dz = \int_0^3 \int_0^{2\pi} \left[\frac{1}{3}(z^2 + R^2)^{3/2}\right]_0^z d\theta dz$$

$$= \frac{1}{3}\int_0^3 \int_0^{2\pi} (2z^2)^{3/2} - (z^2)^{3/2}\, d\theta dz$$

$$= \frac{2\sqrt{2} - 1}{3}\int_0^3 \int_0^{2\pi} z^3 dz$$

$$= 2\pi\, \frac{2\sqrt{2} - 1}{3}\left[\frac{z^4}{4}\right]_0^3$$

$$= 2\pi\, \frac{2\sqrt{2} - 1}{3}\,\frac{81}{4} = \frac{27\pi}{2}(2\sqrt{2} - 1)$$

Example 9.16 Show that if $u = x + y + z$, $uv = y + z$, $uvw = z$ then the Jacobian $\dfrac{\partial(x, y, z)}{\partial(u, v, w)} = u^2v$. Hence evaluate the triple integral $\iiint_V \exp(-(x + y + z)^3)\mathrm{d}V$ where V is the volume enclosed by the planes $x = 0$, $y = 0$, $z = 0$ and $x + y + z = 1$.

Solution Writing x, y and z in terms of u, v and w gives

$$z = uvw,\ y = uv - uvw,\ x = u - uv + uvw$$

so that we can calculate the partial derivatives as follows:

$$\frac{\partial x}{\partial u} = 1 - v + vw, \qquad \frac{\partial x}{\partial v} = -u + uw, \qquad \frac{\partial x}{\partial w} = uv,$$

$$\frac{\partial y}{\partial u} = v - vw, \qquad \frac{\partial y}{\partial v} = u - uw, \qquad \frac{\partial y}{\partial w} = -uv,$$

$$\frac{\partial z}{\partial u} = vw, \qquad \frac{\partial z}{\partial v} = uw, \qquad \frac{\partial z}{\partial w} = uv.$$

Whence the Jacobian is:

$$\begin{aligned}\frac{\partial(x, y, z)}{\partial(u, v, w)} &= \begin{vmatrix} 1 - v + vw & -u + uw & uv \\ v - vw & u - uw & -uv \\ vw & uw & uv \end{vmatrix} \\ &= uv\left(\begin{vmatrix} v - vw & u - uw \\ vw & uw \end{vmatrix} + \begin{vmatrix} 1 - v + vw & -u + uw \\ vw & uw \end{vmatrix} + \begin{vmatrix} 1 - v + vw & -u + uw \\ v - vw & u - vw \end{vmatrix}\right) \\ &= uv(0 + uw - uvw + uvw^2 + uvw - uvw^2 + (u - uw)(1 - v + vw + v - vw)) \\ &= uv(uw + u - uw) \\ &= u^2v \text{ as required. (Computer algebra is particularly good for determinants.)}\end{aligned}$$

We now need to see how the limits transform from $x = 0$, $y = 0$, $z = 0$, $x + y + z = 1$ to u, v, w space. The discerning reader will be able to see that this is potentially quite a tricky problem. The fact that the Jacobian $\dfrac{\partial(x, y, z)}{\partial(u, v, w)} = 0$ where $u = 0$ or $v = 0$ means that the transformation is *singular* if either $u = 0$ or $v = 0$. Both of these conditions impinge on our transformed tetrahedron. The vertices of the tetrahedron in (x, y, z) space are: $(0, 0, 0)$, $(1, 0, 0)$, $(0, 1, 0)$ and $(0, 0, 1)$. The facts that under the transformation $(0, 1, 0) \to (1, 1, 0)$ in (u, v, w) space and that $(0, 0, 1) \to (1, 1, 1)$ in (u, v, w) space are beyond dispute as at these points the mapping is one to one. If all of x, y and z are zero then the only information we have is that $u = 0$, hence the origin can map to any point in the $u = 0$ plane. Also the point $(1, 0, 0)$ in (x, y, z) space can map to any point on the line $u = 1$, $v = 0$. We have therefore deduced that the unit cube $0 \leq u, v, w \leq 1$ in (u, v, w) space is one possible representation of the transformed tetrahedron $x = 0$, $y = 0$, $z = 0$, $x + y + z = 1$ in (x, y, z) space. This is shown schematically in Figure 9.18.

So
$$\begin{aligned}\iiint_V \exp(-(x + y + z)^3)\mathrm{d}V &= \int_0^1\int_0^1\int_0^1 \mathrm{e}^{-u^3}\, u^2v\,\mathrm{d}w\,\mathrm{d}v\,\mathrm{d}u \\ &= \int_0^1 \left[\tfrac{1}{2}\, v^2u^2\mathrm{e}^{-u^3}\right]_0^1 \mathrm{d}u \\ &= \tfrac{1}{2}\left[-\tfrac{1}{3}\,\mathrm{e}^{-u^3}\right]_0^1 \\ &= \frac{1}{6}(1 - \mathrm{e}^{-1}), \text{ by elementary integration.}\end{aligned}$$

Example 9.17 Determine the moment of inertia of a uniform solid hemisphere of radius a about any diameter.

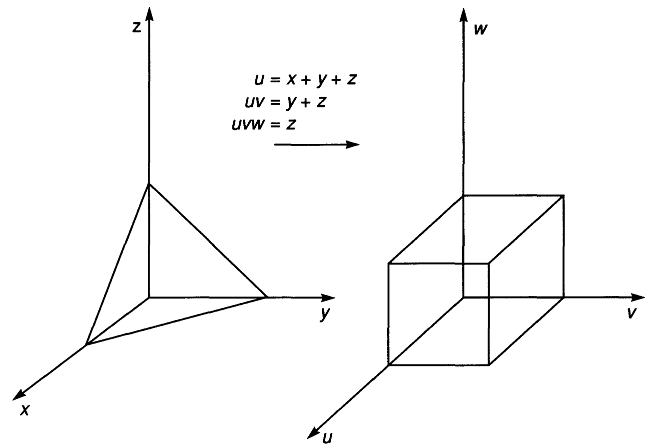


Figure 9.18 A schematic picture of the transformation of V from (x, y, z) co-ordinates to (u, v, w) co-ordinates.

Solution The moment of inertia of a rigid body is defined as the quantity

$$I = \int_V \rho d^2 \mathrm{d}V$$

where ρ is the density of the body and d is the perpendicular distance of a particle of the body from the axis of rotation (see the author's *Work Out Mechanics*, page 127 for more on this). Here we are less concerned about the physical mechanics than with the 'mechanics' of evaluating the integral for I. Take the z-axis to be the axis about which the moment of inertia is to be taken. Figure 9.19 shows the spherical polar co-ordinates we use.

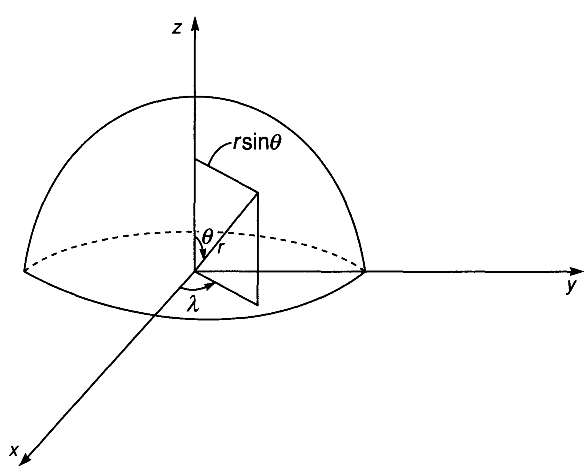


Figure 9.19 The hemisphere and (r, θ, λ) co-ordinate system.

The perpendicular distance of an arbitrary point of the hemisphere from the z-axis is therefore $r \sin\theta$ (and not r which is a common error). Also, the element of volume in spherical polars (r,θ,λ) is $r^2 \sin\theta \mathrm{d}r\mathrm{d}\theta\mathrm{d}\lambda$. Thus the moment of inertia becomes the triple integral

$$I = \int_V \rho d^2 \, \mathrm{d}V = \int_0^a \int_0^\pi \int_0^\pi (r \sin\theta)^2 r^2 \sin\theta \mathrm{d}\theta \mathrm{d}\lambda \mathrm{d}r$$

This integration is reasonably straightforward, either using software or as follows:

$$\begin{aligned}\int_0^a \int_0^\pi \int_0^\pi \rho (r \sin\theta)^2 r^2 \sin\theta \mathrm{d}\theta \mathrm{d}\lambda \mathrm{d}r &= \int_0^a \int_0^\pi \left[-\cos\theta + \frac{1}{3}\cos^3\theta \right]_0^\pi \rho r^4 \, \mathrm{d}\lambda \mathrm{d}r \\ &= \int_0^a \frac{4}{3} \rho r^4 \cdot \pi \mathrm{d}r \\ &= \frac{4}{15} \rho a^5 \pi\end{aligned}$$

It is usual to express this as an intergal involving the mass of the hemisphere which is $m = \frac{2}{3}\rho\pi a^3$, whence $I = \frac{4}{15}a^5\pi \cdot \frac{3m}{2a^3\pi} = \frac{2}{5}ma^2$. This is the same as the moment of inertia of a solid *sphere* about any diameter.

Example 9.18

The nose cone of a rocket has the shape of the paraboloid $z = a^2 - x^2 - y^2$. Calculate the moment of inertia about its axis of symmetry in terms of its mass m and the base radius a assuming it is of uniform density.

Solution

This problem is best tackled in cylindrical polar co-ordinates (R, θ, z) in which the nose cone has equation $z = a^2 - R^2$. The co-ordinates are shown in Figure 9.20.

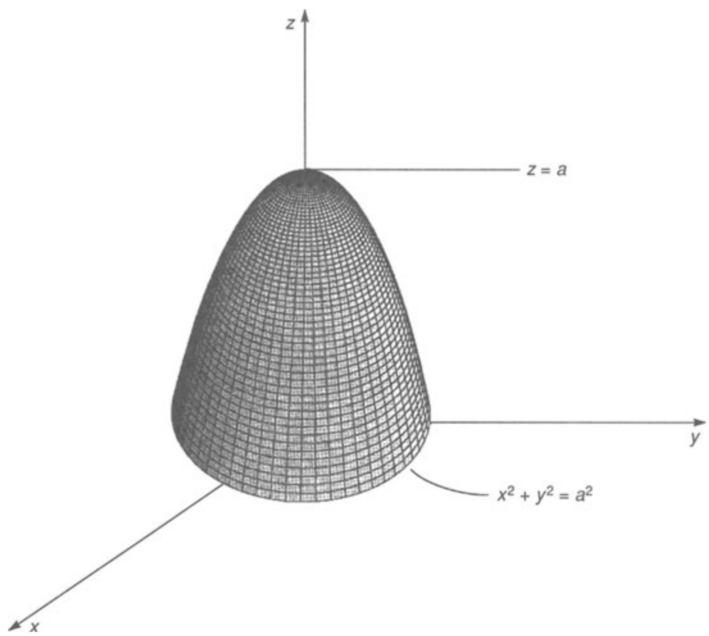


Figure 9.20 The nose cone (paraboloid) $z = a^2 - x^2 - y^2$ (≥ 0).

The moment of inertia is given by $I = \int_V \rho R^2 \mathrm{d}V$ where ρ is the density. Cylindrical polars are natural for moments of inertia as R itself is the perpendicular distance to the z-axis. Hence

$$
\begin{aligned}
I &= \int_0^{a^2}\int_0^{\sqrt{a^2-z}}\int_0^{2\pi} \rho R^2 \cdot R\mathrm{d}\theta\mathrm{d}R\mathrm{d}z \\
&= 2\pi \int_0^{a^2} \left[\rho \frac{1}{4} R^4\right]_0^{\sqrt{a^2-z}} \mathrm{d}z \\
&= \frac{\pi\rho}{2}\int_0^{a^2} (a^2 - z)^2\,\mathrm{d}z \\
&= \frac{\pi\rho}{2}\left[a^4 z - a^2 z^2 + \frac{1}{3}z^3\right]_0^{a^2} \\
&= \frac{1}{6}\pi\rho\, a^6
\end{aligned}
$$

We now need to calculate the mass as we do not know this in the same way as we knew the mass (volume) of a hemisphere. The mass of the nose cone is its density times its volume, whence

$$
\begin{aligned}
m &= \rho \int_0^{a^2}\int_0^{\sqrt{a^2-z}}\int_0^{2\pi} R\mathrm{d}\theta\mathrm{d}R\mathrm{d}z \\
&= 2\pi\rho \int_0^{a^2} \left[\frac{1}{2} R^2\right]_0^{\sqrt{a^2-z}} \mathrm{d}z
\end{aligned}
$$

$$= \pi\rho \int_0^{a^2} (a^2 - z)\mathrm{d}z$$

$$= \pi\rho\left[a^2 z - \tfrac{1}{2} z^2\right]_0^{a^2} = \frac{\pi}{2} \rho a^4$$

Hence, eliminating the density ρ we have

$$I = \frac{1}{6} \pi a^6 \cdot \frac{2m}{\pi a^4} = \frac{1}{3} ma^2$$

9.3 Exercises

9.1. Evaluate the following double integrals in the x–y plane:

(a) $\int_D\int x^2 y\mathrm{d}x\mathrm{d}y$ where D is the rectangle bounded by the lines $x = 0, \quad x = 1, \quad y = 0, \quad y = 2.$

(b) $\int_D\int (x^2 + y^2)\mathrm{d}x\mathrm{d}y$ where D is the triangle with vertices (0,0), (0,1) and (1,0).

(c) $\int_D\int \sin(x + y)\mathrm{d}x\mathrm{d}y$ where D is the square $0 \le x, y \le 1$, hence show that $0 \le \int_D\int \sin(x + y)\mathrm{d}x\mathrm{d}y \le 1$.

(d) $\int_D\int (3xy^2 - y)\mathrm{d}x\mathrm{d}y$ where D is the region between $y = |x|$, and $y = -|x|$ where $-1 \le x \le 1$.

9.2. Draw the domain of and hence evaluate the following double integrals:

(a) $\int_D\int \sqrt{xy}\mathrm{d}x\mathrm{d}y$ where D is the domain $0 \le y \le 1$, $y^2 \le x \le y$.

(b) $\int_D\int (x^4 + y^2)\mathrm{d}x\mathrm{d}y$ where D is the region between the curves $y = x^3$ and $y = x^2$.

(c) $\int_D\int e^{x^2}\,\mathrm{d}x\mathrm{d}y$ where D is the triangle formed by the x-axis, $2y = x$ and $x = 2$.

(d) $\int_D\int (x^{1/2} - y^2)\mathrm{d}x\mathrm{d}y$ where D is the area between $y = x^2$ and $y = x^{1/4}$.

9.3. For the following integrals, draw the domain and hence find the new limits if you reverse the order of integration:

(a) $\int_0^1\int_{x^4}^{x^2} f(x, y)\mathrm{d}y\mathrm{d}x$, (b) $\int_0^1\int_0^{y^2} f(x, y)\mathrm{d}x\mathrm{d}y$,

(c) $\int_0^1\int_{-y}^{y} f(x, y)\mathrm{d}x\mathrm{d}y$, (d) $\int_1^3\int_{-x}^{x^2} f(x, y)\mathrm{d}y\mathrm{d}x$.

9.4. Sketch the domain of integration and hence evaluate the following:

(a) $\int_1^2\int_x^{2x} \frac{\mathrm{d}y\mathrm{d}x}{(x^2 + y^2)}$, (b) $\int_0^1\int_x^{\sqrt{2-x^2}} \frac{x}{\sqrt{x^2 + y^2}}\mathrm{d}y\mathrm{d}x$.

9.5. Evaluate the following integrals by reversing the order of integration:

(a) $\int_0^1\int_{\sqrt{x}}^1 \sin\left(\frac{y^3 + 1}{2}\right)\mathrm{d}y\mathrm{d}x$, (b) $\int_0^1\int_{x^2}^1 \frac{x^3}{\sqrt{x^4 + y^2}}\,\mathrm{d}y\mathrm{d}x$

(c) $\int_0^1\int_0^{\cos^{-1}y} e^{\sin x}\mathrm{d}x\mathrm{d}y$, (d) $\int_0^1\int_x^1 x^2 e^{y^4}\,\mathrm{d}y\mathrm{d}x$.

9.6. The curve $R = 3\cos 2\theta$ represents a *lemniscate* or *two-leafed rose* in plane polar co-ordinates (R, θ). Write down its area as a double integral and hence evaluate it. (See Chapter 11 for another approach to finding areas.)

9.7. Evaluate the double integral $\int\int_A x^2y^2\mathrm{d}A$ where A is the disc $x^2 + y^2 \le 1$, $z = 0$.

9.8. Use plane polar co-ordinates to evaluate the integral

$$\int\int_D \frac{x^2 - y^2}{x^2 + y^2}\,\mathrm{d}x\mathrm{d}y$$

where D is that part of the x–y plane bounded by the parabola $y^2 = 4(1 - x)$ and the co-ordinate axes in the positive quadrant.

9.9. By transforming to plane polar co-ordinates, evaluate the double integral

$$\int_0^{\frac{1}{2}\sqrt{2}}\int_x^{\sqrt{1-x^2}} \frac{\ln(x^2 + y^2)}{\sqrt{x^2 + y^2}}\,\mathrm{d}y\mathrm{d}x$$

9.10. Use the transformation indicated to evaluate the following integrals:

(a) $\int\int_D x^3(1 - x^4 - y^4)\mathrm{d}x\mathrm{d}y$ where D is the region $x \ge 0, y \ge 0$, $x^4 + y^4 \le 1$, and $x = \sqrt{r\cos\theta}$, $y = \sqrt{r\sin\theta}$.

(b) $\int\int_D \ln(x^2 + y^2)\mathrm{d}x\mathrm{d}y$ where D is the region of the annulus $a^2 \le (x^2 + y^2) \le b^2$ in the first quadrant. Use polar co-ordinates.

(c) $\int_0^1\int_y^{2-y} \frac{x + y}{x^2}\, e^{x+y}\,\mathrm{d}x\mathrm{d}y$ using the transformation $x + y = u, \frac{y}{x} = v.$

9.11. Use Green's Theorem in the Plane to evaluate the integral

$$\oint_C (1 + 10xy + y^2)\mathrm{d}x + (6xy + 5x^2)\mathrm{d}y$$

where C is the square with corners (0,0), $(a,0)$, $(0,a)$ and (a,a).

9.12. Use Green's Theorem to show that

$$\oint_C e^x \sin y\mathrm{d}x + e^x \cos y\mathrm{d}y = 0$$

for any closed contour C. Interpret this in terms of potentials.

9.13. Evaluate $\int_V e^{-r^2}\,dV$ where V is the whole of space ($r^2 = x^2 + y^2 + z^2$).

9.14. Evaluate the following triple integrals:

(a) $\int_0^1\int_0^2\int_1^3 x^2yz\,dzdydx$, (b) $\int_0^2\int_1^3\int_2^4 xy^2z\,dzdydx$,

(c) $\int_0^1\int_{x^2}^{\sqrt{x}}\int_{x-z}^{x+z} (x + y + z)\,dydzdx$.

9.15. Sketch the region contained between the parabolic cylinders $y = x^2$ and the planes $z = 0$ and $x + y + z = 2$. Show that the volume can be written as the triple integral

$$\int_0^1\int_{x^2}^{\sqrt{x}}\int_0^{2-x-y} dzdydx$$

9.16. Find the volume of the solid bounded by the planes $z = 0$, $y = 0$, $x = 0$, the cylinder $x^2 + y^2 = 4$ and the hyperbolic paraboloid $z = 6 - xy$ in the positive octant.

9.17. Find the volume of the region bounded above by the sphere $x^2 + y^2 + z^2 = a^2$ and below by the cone $z^2\sin^2\alpha = (x^2 + y^2)\cos^2\alpha$, where α is a constant such that $0 \le \alpha \le \pi$. Use your result to confirm that a sphere of radius a has a volume $\frac{4}{3}\pi a^3$.

9.18. Evaluate the triple integral

$$\iiint_V \sqrt{\frac{x}{y^2 + z^2}}\,dV$$

where V is the volume bounded by the cone $z^2 + y^2 = x^2$, the cylinder $y^2 + z^2 = 4$ and the planes $x = 0$ and $x = 2$.

9.19. Using two-dimensional polar co-ordinates (R, θ) prove that the co-ordinates of the centre of mass (in *Cartesian* co-ordinates) $(\bar{x}, \bar{y})$ of an area D are given by the formulae:

$$\bar{x} = \frac{1}{\text{Area}}\iint_D R^2\cos\theta\, dRd\theta$$

$$\bar{y} = \frac{1}{\text{Area}}\iint_D R^2\cos\theta\, dRd\theta$$

Use these formulae to find the co-ordinates of the centre of mass of the following laminas that have constant thicknesses:
(a) the cardioid $R = a(1 + \cos\theta)$,
(b) the quarter lemniscate $R^2 = 2a^2\cos 2\theta$, $0 \le \theta \le \pi/2$.

9.20. Find the surface area of that part of the sphere $x^2 + y^2 + z^2 = 4$ that lies outside the cylinder $x^2 + y^2 = 1$.

9.21. The formulae $x = ar(\cos\theta)^\alpha$, $y = br(\sin\theta)^\alpha$ define generalised polar co-ordinates (r, θ). Determine the Jacobian of this transformation.

9.22. Evaluate the integral $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{e^{-(x-y)^2}}{1 + (x + y)^2}\,dxdy$ by integrating over the square $-a \le x \le a$, $-a \le y \le a$ and taking the limit as $a \to \infty$.

(*Hint:* use the transformation $u = x - y$, $v = x + y$ and see Example 9.11.)

9.23. Use triple integration to determine the moment of inertia of a cone of height h and base radius a about its axis of symmetry.

9.24. Find the moment of inertia of the ice cream cone bounded below by the half cone $z = \sqrt{3(x^2 + y^2)}$ and bounded above by the unit sphere $x^2 + y^2 + z^2 = 1$, about the axis of symmetry.

Topic Guide

Definition
Evaluation by Projection
and by Parameterisation

10 Surface Integrals

10.1 Fact Sheet

This chapter deals with integrals over arbitrarily shaped surfaces. A surface is usually denoted by the letter S. We are only going to consider surfaces with two sides, that is we ignore Möbius Strips, Klein Bottles and like topological oddities. S therefore has two sides; if it is closed, an inside and an outside. A normal, usually a unit normal written $\hat{\mathbf{n}}$, is deemed positive if it is drawn away from the (positive) outside of a surface. In general, the positive side of a surface is that which has positive curvature; for a flat surface, either side may be chosen. An element of area, dS, is doubly infinitesimal and the directed element of area $d\mathbf{S} = \hat{\mathbf{n}}dS$. A surface integral is written as either

$$\int_S \mathbf{A} \cdot d\mathbf{S} \qquad \text{or} \qquad \int_S\int \mathbf{A} \cdot d\mathbf{S}$$

where $\mathbf{A}$ is a continuous vector field. Other surface integrals are $\int_S \phi d\mathbf{S}$, $\int_S \mathbf{A} \times d\mathbf{S}$, $\int_S \phi dS$.

The first two are vectors, the third is a scalar. Older books use the notation $\oiint_S$ for the integral over a closed surface S. To evaluate a surface integral, the object is to turn it into a double integral of the type met in Chapter 9. There are two distinct ways of doing this, the first is to use *projection* whereby the curved surface S is projected on to one of the co-ordinate planes (usually the x–y plane). If the shadow of S on the x–y plane is called R then the formula for calculating $\int_S \mathbf{A} \cdot d\mathbf{S}$ is $\int_S \mathbf{A} \cdot d\mathbf{S} = \int_R \mathbf{A} \cdot \hat{\mathbf{n}} \dfrac{dxdy}{|\hat{\mathbf{n}} \cdot \hat{\mathbf{k}}|}$. Obviously, $|\hat{\mathbf{n}} \cdot \hat{\mathbf{k}}| \neq 0$ here. The second way of calculating surface integrals is by using parameterisation. This method involves defining parameters u, v such that $\int_S \ldots dS$ is transformed directly into a double integral over u and v of the form $\iint \ldots dudv$, an integral in parameter space that is evaluated in the normal way. The integral $\int_S dS$ is the area of the curved surface S (see Example 10.5).

10.2 Worked Examples

Example 10.1 An open surface S is such that the unit normal $\hat{\mathbf{n}}$ at any point on it is never at right angles to the z direction $\hat{\mathbf{k}}$. Let R be the projection of S on to the x–y plane, and show that if $\mathbf{A}$ is a vector field that is continuous on S then

$$\int_S \mathbf{A} \cdot d\mathbf{S} = \int_R \mathbf{A} \cdot \hat{\mathbf{n}} \frac{dxdy}{|\hat{\mathbf{n}}.\hat{\mathbf{k}}|}$$

Solution

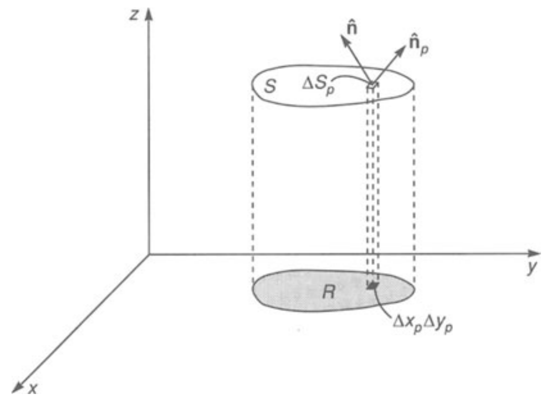


Figure 10.1 **The surface S and its projection R on to the x–y plane.**

The situation is displayed in Figure 10.1.

The definition of a surface integral using the Riemann formulation is in terms of the limit of a sum $\int_S \mathbf{A} \cdot \mathrm{d}\mathbf{S} = \int_S \mathbf{A} \cdot \hat{\mathbf{n}}\mathrm{d}S = \lim_{\substack{\Delta S_p \to 0 \\ M\to\infty}} \sum_{p=1}^{M} \mathbf{A}_p \cdot \hat{\mathbf{n}}_p \Delta S_p$ where $\Delta S_p \to 0$ in such a way that $\max(\Delta S_p) \to 0$.

Now $|\hat{\mathbf{n}}_p \cdot \hat{\mathbf{k}}| = \cos\gamma$ where γ is the angle between $\hat{\mathbf{n}}$ and $\hat{\mathbf{k}}$, hence it is possible to write $|\Delta\mathbf{S}_p \cdot \hat{\mathbf{k}}| = |\hat{\mathbf{n}}\Delta S_p \cdot \hat{\mathbf{k}}| = \Delta S_p \cos\gamma = \Delta x \Delta y$ as the projection of the infinitesimal element of the surface S, ΔS_p, on to the x–y plane. Thus we may write $|\Delta\mathbf{S}_p \cdot \hat{\mathbf{k}}| = \Delta x \Delta y$

or
$$\Delta S_p = \frac{\Delta x \Delta y}{|\hat{\mathbf{n}}_p \cdot \mathbf{k}|}$$

Thus
$$\int_S \mathbf{A} \cdot \mathrm{d}\mathbf{S} = \lim_{\substack{\Delta S_p \to 0 \\ M\to\infty}} \sum_{p=1}^{M} \mathbf{A}_p \cdot \hat{\mathbf{n}}_p \frac{\Delta x \Delta y}{|\hat{\mathbf{n}}_p \,.\, \hat{\mathbf{k}}|}$$

$$= \int_R \mathbf{A} \cdot \mathbf{n} \frac{\mathrm{d}x\mathrm{d}y}{|\mathbf{n}_p \cdot \mathbf{k}|}$$ omitting the carats, and where R is the region of the x–y plane corresponding to the projection of S. This establishes the result.

Example 10.2

Evaluate the integral $\int_S A\mathrm{d}S$ where $A = \dfrac{z}{(4x^2 + 4y^2 + 1)^{1/2}}$ and S is the surface of the paraboloid $z = 4 - x^2 - y^2$, $0 \leq z \leq 4$.

Solution

It is better to use the projection method here, projecting the paraboloid on to the x–y plane as follows:

$$\int_S A\mathrm{d}S = \int_R A \frac{\mathrm{d}R}{|\hat{\mathbf{n}} \cdot \hat{\mathbf{k}}|}$$

where $\hat{\mathbf{n}}$ is the unit normal to the surface S, and is found by using the gradient operator in the following way: $\nabla(z + x^2 + y^2 - 4) = 2x\mathbf{i} + 2y\mathbf{j} + \mathbf{k}$

so $\hat{\mathbf{n}} = \dfrac{\nabla(z + x^2 + y^2 - 4)}{|\nabla(z + x^2 + y^2 - 4)|} = \dfrac{2x\mathbf{i} + 2y\mathbf{j} + \mathbf{k}}{(4x^2 + 4y^2 + 1)^{1/2}}$

Therefore $\hat{\mathbf{n}} \cdot \hat{\mathbf{k}} = (4x^2 + 4y^2 + 1)^{-1/2}$ and so $\dfrac{A}{\hat{\mathbf{n}} \cdot \hat{\mathbf{k}}} = \dfrac{z}{(4x^2 + 4y^2 + 1)^{1/2}} (4x^2 + 4y^2 + 1)^{1/2}$

$= z = 4 - x^2 - y^2$

on S. Hence $\int_R A \dfrac{\mathrm{d}R}{|\hat{\mathbf{n}} \cdot \hat{\mathbf{k}}|} = \int_R (4 - x^2 - y^2)\mathrm{d}R$

$$= \int_0^{2\pi}\int_0^2 (4 - \rho^2)\rho\mathrm{d}\rho\mathrm{d}\theta$$ in cylindrical polar co-ordinates,

where we have had to use the symbol ρ instead of the more normal R for the radius in order to avoid R having two meanings. S and R are shown in Figure 10.2.

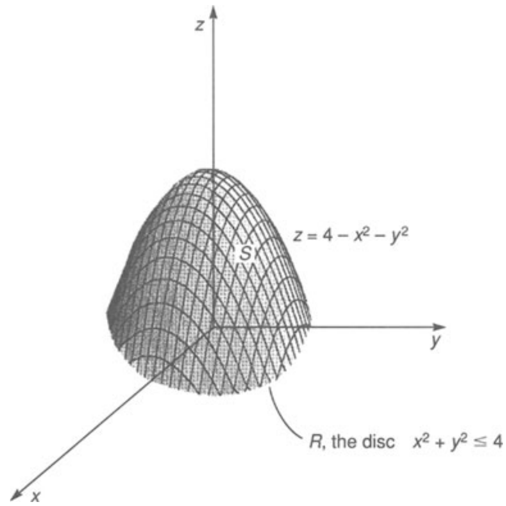


Figure 10.2 The paraboloid surface $z = 4 - x^2 - y^2$.

Thus we evaluate the double integral, which is straightforward:

$$\int_S A\mathrm{d}S = \int_0^{2\pi} \left[2\rho^2 - \tfrac{1}{4}\rho^4\right]_0^2 \mathrm{d}\theta$$

$$= \int_0^{2\pi} 8 - \tfrac{16}{4}\mathrm{d}\theta = 2\pi \times 4 = 8\pi$$

Example 10.3

If $\mathbf{F} = x\mathbf{i} - y\mathbf{j} + z^2\mathbf{k}$ find $\int_0 \mathbf{F} \cdot \hat{\mathbf{n}}\mathrm{d}S$ where S is that part of the plane $x + y + z = 1$ which lies in the first quadrant.

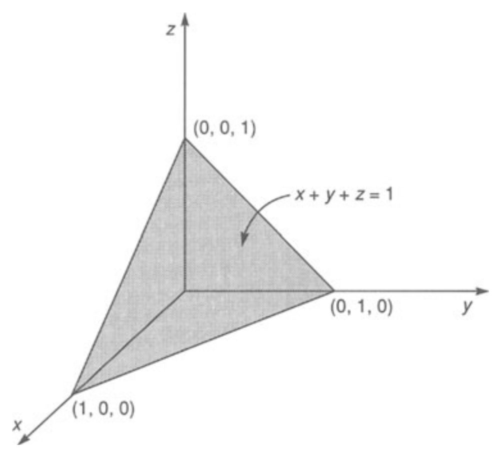


Figure 10.3 The plane $x + y + z = 1$.

Solution The plane is shown in Figure 10.3. The normal to S is given by $\nabla(x + y + z - 1) = \mathbf{i} + \mathbf{j} + \mathbf{k}$ so the unit normal is given by $\frac{1}{\sqrt{3}}(\mathbf{i} + \mathbf{j} + \mathbf{k})$. Also, $\mathbf{F} \cdot \hat{\mathbf{n}} = \frac{1}{\sqrt{3}}(x - y + z^2) = \frac{1}{\sqrt{3}}(x - y + (1 - x - y)^2)$ on S. Projecting on to the x–y plane, $\hat{\mathbf{n}} \cdot \hat{\mathbf{k}} = \frac{1}{\sqrt{3}}$ so $\int_S \mathbf{F} \cdot \hat{\mathbf{n}}\mathrm{d}S = \int_R \mathbf{F} \cdot \hat{\mathbf{n}} \frac{\mathrm{d}x\mathrm{d}y}{|\hat{\mathbf{n}} . \hat{\mathbf{k}}|}$ implies $\int_S \mathbf{F} \cdot \hat{\mathbf{n}}\mathrm{d}S = \int_0^1\int_0^{1-x} x - y + (1 - x - y)^2\mathrm{d}y\mathrm{d}x$

$$= \int_0^1\int_0^{1-x} 1 - x - 3y + 2xy + x^2 + y^2\mathrm{d}y\mathrm{d}x$$

$$= \int_0^1 \left[y - xy - \tfrac{3}{2}y^2 + xy^2 + x^2y + \tfrac{1}{3}y^3\right]_0^{1-x} \mathrm{d}x$$

$$= \int_0^1 1 - x - x(1 - x) - \tfrac{3}{2}(1 - x)^2 + x(1 - x)^2 + x^2(1 - x) + \tfrac{1}{3}(1 - x)^3\,\mathrm{d}x$$

This integral is a little long, but it is straightforward. The integrand simplifies a little to $(1 - x)(-\frac{1}{2}(1 - x) + x(1 - x) + x^2 + \frac{1}{3}(1 - x)^2)$, so putting $u = 1 - x$ also helps, and the integral becomes $\int_0^1 (-\frac{3}{2}u^2 + u + \frac{1}{3}u^3)du = \left[-\frac{1}{2}u^3 + \frac{1}{2}u^2 + \frac{1}{12}u^4\right]_0^1 = \frac{1}{12}$.

Example 10.4 Evaluate the surface integral $\int_S \mathbf{F} \times d\mathbf{S}$ where $\mathbf{F} = z\hat{\mathbf{k}}$ and S is that portion of the unit sphere $x^2 + y^2 + z^2 = 1$ that is in the positive octant $x, y, z \geq 0$.

Solution We shall use projection to evaluate this integral. Project S on to the x–y plane and use the relationship $\int_S \mathbf{F} \times d\mathbf{S} = \int_R \mathbf{F} \times \hat{\mathbf{n}} \frac{dxdy}{|\hat{\mathbf{n}} \cdot \hat{\mathbf{k}}|}$. Region R is shown in Figure 10.4.

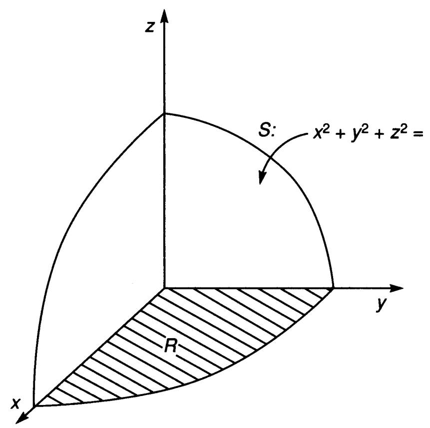


Figure 10.4 The eighth sphere S and its projection on to the x–y plane R.

The unit normal $\hat{\mathbf{n}}$ to the surface S obeys $d\mathbf{S} = \hat{\mathbf{n}}dS$ and is, for the unit sphere centre the origin, the position vector. That is $\hat{\mathbf{n}} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ and so $|\hat{\mathbf{n}} \cdot \hat{\mathbf{k}}| = (x\mathbf{i} + y\mathbf{j} + z\mathbf{k}) \cdot \mathbf{k} = z$. Now $\mathbf{F} = z\hat{\mathbf{k}}$ so $\mathbf{F} \times \hat{\mathbf{n}} = z\hat{\mathbf{k}} \times (x\mathbf{i} + y\mathbf{j} + z\mathbf{k}) = z(x\mathbf{j} - y\mathbf{i})$. Thus $\int_S \mathbf{F} \times d\mathbf{S} = \int_R z(x\mathbf{j} - y\mathbf{i}) \frac{dxdy}{z}$ from which the z cancels, saving us the trouble of expressing it in terms of x and y via the algebraic form of S. The region R is the quarter circle $x^2 + y^2 \leq 1$, $x, y \geq 0$ (see Figure 10.5), thus the required integral is

$$\begin{aligned}\int_R (x\mathbf{j} - y\mathbf{i})dxdy &= \int_0^1 \int_0^{\sqrt{1-x^2}} (x\mathbf{j} - y\mathbf{i})dydx \\ &= \int_0^1 x\sqrt{1 - x^2}\,\mathbf{j} - \tfrac{1}{2}(1 - x^2)\mathbf{i}dx \\ &= \frac{1}{3}\mathbf{j} - \frac{1}{3}\mathbf{i}\end{aligned}$$

omitting the gory details. (The first integral requires the substitution $u = 1 - x^2$, the second is straightforward.) Note that the answer is a vector as would be expected from the cross product in

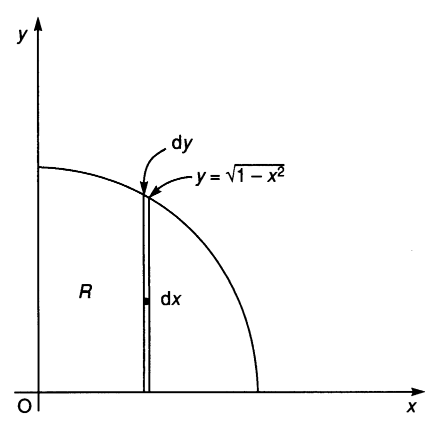


Figure 10.5 Integrating through the region R.

$\int_S \mathbf{F} \times d\mathbf{S}$. Try this problem yourself using plane polar co-ordinates, and you will find it a little easier (I hope).

Example 10.5 A surface S is described by the parameterisation $\mathbf{r} = \mathbf{r}(\theta, \phi)$. Show that the normal to S is given by $\dfrac{\partial \mathbf{r}}{\partial \theta} \times \dfrac{\partial \mathbf{r}}{\partial \phi}$. Deduce the surface area of the torus given parametrically by
$\mathbf{r} = (a + b\cos\phi)\cos\theta\,\mathbf{i} + (a + b\cos\phi)\sin\theta\,\mathbf{j} + b\sin\phi\,\mathbf{k}$.

Solution On the surface $\mathbf{r} = \mathbf{r}(\theta, \phi)$, if either θ or ϕ is held constant, the resulting curve is embedded in the surface S. To be specific, the curves $\mathbf{r} = \mathbf{r}(\theta_0, \phi)$, $\mathbf{r} = \mathbf{r}(\theta, \phi_0)$ where $\theta = \theta_0$, $\phi = \phi_0$ are embedded in S. The lines $\dfrac{\partial \mathbf{r}}{\partial \phi}$ and $\dfrac{\partial \mathbf{r}}{\partial \theta}$ are thus tangents to the lines $\mathbf{r} = \mathbf{r}(\theta_0, \phi)$ and $\mathbf{r} = \mathbf{r}(\theta, \phi_0)$ respectively. They are *de facto* also tangent lines to the surface S. The plane containing the two lines $\dfrac{\partial \mathbf{r}}{\partial \theta}$ and $\dfrac{\partial \mathbf{r}}{\partial \phi}$ is thus a tangent plane to S that touches S at the point $\theta = \theta_0$, $\phi = \phi_0$. The vector $\dfrac{\partial \mathbf{r}}{\partial \theta} \times \dfrac{\partial \mathbf{r}}{\partial \phi}$ is normal to both $\dfrac{\partial \mathbf{r}}{\partial \theta}$ and $\dfrac{\partial \mathbf{r}}{\partial \phi}$ (if this is not obvious from the definition of cross product, note that the scalar triple products $\left(\dfrac{\partial \mathbf{r}}{\partial \theta} \times \dfrac{\partial \mathbf{r}}{\partial \phi}\right) \cdot \dfrac{\partial \mathbf{r}}{\partial \theta}$ and $\left(\dfrac{\partial \mathbf{r}}{\partial \theta} \times \dfrac{\partial \mathbf{r}}{\partial \phi}\right) \cdot \dfrac{\partial \mathbf{r}}{\partial \phi}$ are both zero). As shown in Figure 10.6, the vector element of area

$$d\mathbf{S} = \frac{\partial \mathbf{r}}{\partial \theta} d\theta \times \frac{\partial \mathbf{r}}{\partial \phi} d\phi = \left(\frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \phi}\right) d\theta d\phi$$

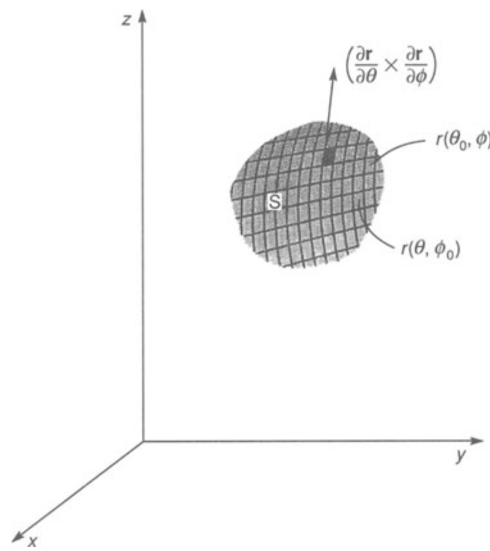


Figure 10.6 The surface S, showing the curves $\mathbf{r}(\theta_0, \phi)$, $\mathbf{r}(\theta, \phi_0)$ and the vector $\left(\dfrac{\partial \mathbf{r}}{\partial \theta} \times \dfrac{\partial \mathbf{r}}{\partial \phi}\right)$ normal to an element of S.

Thus

$$dS = \hat{\mathbf{n}} \cdot d\mathbf{S} = \frac{\left(\dfrac{\partial \mathbf{r}}{\partial \theta} \times \dfrac{\partial \mathbf{r}}{\partial \phi}\right)}{\left|\left(\dfrac{\partial \mathbf{r}}{\partial \theta} \times \dfrac{\partial \mathbf{r}}{\partial \phi}\right)\right|} \cdot \left(\frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \phi}\right) d\theta d\phi$$

$$= \left|\left(\frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \phi}\right)\right| d\theta d\phi$$

We have thus derived the potentially useful formula

$$\int_S dS = \iint \left|\left(\frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \phi}\right)\right| d\theta d\phi$$

for calculating the area of a curved surface. The question now requests us to apply this to finding the surface area of a torus, which is shown in Figure 10.7.

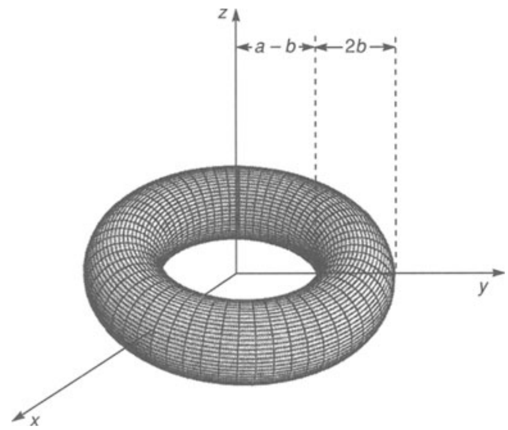


Figure 10.7 The torus.

With $\mathbf{r} = (a + b\cos\phi)\cos\theta\,\mathbf{i} + (a + b\cos\phi)\sin\theta\,\mathbf{j} + b\sin\phi\,\mathbf{k}$

$$\frac{\partial \mathbf{r}}{\partial \theta} = -(a + b\cos\phi)\sin\theta\,\mathbf{i} + (a + b\cos\phi)\cos\theta\,\mathbf{j}$$

and $$\frac{\partial \mathbf{r}}{\partial \phi} = -b\sin\phi\cos\theta\,\mathbf{i} - b\sin\phi\sin\theta\,\mathbf{j} + b\cos\phi\,\mathbf{k}$$

so

$$\begin{aligned}\frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \theta} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -(a + b\cos\phi)\sin\theta & (a + b\cos\phi)\cos\theta & 0 \\ -b\sin\phi\cos\theta & -b\sin\phi\sin\theta & b\cos\phi \end{vmatrix} \\ &= (a + b\cos\phi)b\cos\theta\cos\phi\,\mathbf{i} + (a + b\cos\phi)b\sin\theta\cos\phi\mathbf{j} \\ &\quad + (a + b\cos\phi)b\sin\phi\,\mathbf{k}\end{aligned}$$

So

$$\begin{aligned}\left|\frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \phi}\right|^2 &= b^2(a + b\cos\phi)^2(\cos^2\theta\cos^2\phi + \sin^2\theta\cos^2\phi + \sin^2\phi) \\ &= b^2(a + b\cos\phi)^2\end{aligned}$$

Thus we obtain the simple expression $\mathrm{d}S = b(a + b\cos\phi)\mathrm{d}\theta\mathrm{d}\phi$. The surface area of the torus is therefore

$$\begin{aligned}\int_S \mathrm{d}S &= \int_0^{2\pi}\int_0^{2\pi} b(a + b\cos\phi)\mathrm{d}\theta\mathrm{d}\phi \\ &= 2\pi b\int_0^{2\pi}(a + b\cos\phi)\mathrm{d}\phi \\ &= 2\pi b \cdot 2\pi a = 4\pi^2 ab\end{aligned}$$

Note that this is in fact the area of a rectangle of sides $2\pi a$ and $2\pi b$. A little thought should indicate an interpretation of this. The torus has the shape of an inner tube of a tyre, and it is as if it has been made out of a rectangle of rubber. If this rubber sheet has dimensions $2\pi a \times 2\pi b$ then the inner tube has been manufactured with no distortion. Indeed no distortion does arise out of joining opposite edges of the rectangle to form a cylinder of length $2\pi a$ and radius b. However exactly complementary stretch and shrinkage has taken place as the cylinder is bent to make the torus. This probably does not happen in the manufacture of real inner tubes!

Example 10.6 Evaluate the surface integral $\int_S \mathbf{F} \cdot \mathrm{d}\mathbf{S}$ where $\mathbf{F} = 3y\mathbf{i} + 2x^2\mathbf{j} + z^3\mathbf{k}$ and S is the surface of the cylinder $x^2 + y^2 = 1, \quad 0 < z < 1$.

Solution This problem is best solved using direct parameterisation of the surface S as there is no convenient projection possible.

On S, $x = \cos\theta$, $y = \sin\theta$, $z = z$ $(R = 1)$ and so $dS = 1 \cdot d\theta dz$, the rectangular element dS having sides $1 \times d\theta$ and dz.

Now, $\hat{\mathbf{n}} = \dfrac{\nabla(x^2 + y^2 - 1)}{|\nabla(x^2 + y^2 - 1)|} = x\mathbf{i} + y\mathbf{j}$ on $x^2 + y^2 = 1$ so that

$$\begin{aligned}\mathbf{F} \cdot d\mathbf{S} &= (3yx + 2x^2y)d\theta dz \\ &= (3\cos\theta\sin\theta + 2\cos^2\theta\sin\theta)d\theta dz \qquad (\text{on } S)\end{aligned}$$

Hence

$$\begin{aligned}\int_S \mathbf{F} \cdot dS &= \int_0^{2\pi}\int_0^1 3\cos\theta\sin\theta + 2\cos^2\theta\sin\theta dz d\theta \\ &= \int_0^{2\pi}(3\cos\theta\sin\theta + 2\cos^2\theta\sin\theta)d\theta \\ &= \left[\frac{3}{2}\sin^2\theta - \frac{2}{3}\cos^3\theta\right]_0^{2\pi} = 0\end{aligned}$$

Could we have forseen this zero result? Possibly. Integrals that go all the way around cylinders or spheres always have a good chance of being zero by a continuity argument that in fact applies to all closed surfaces: what goes in must come out. This is only false if there are source or sink terms inside S. The next chapter makes use of these kind of arguments and puts them on a mathematical footing.

Example 10.7 Evaluate the integral $\int_S \mathbf{F} \cdot d\mathbf{S}$ where $\mathbf{F}$ is the vector field $\mathbf{F} = x^2\mathbf{i} + y^2\mathbf{j} + z^2\mathbf{k}$ and S is the closed unit sphere $x^2 + y^2 + z^2 = 1$ (see Exercise 11.3).

Solution As with Example 10.6, there is no convenient projection surface, hence we have no choice but to evaluate this surface integral directly by using parameterisation. On the unit sphere centre the origin we have

$$x = \sin\theta\cos\lambda, \quad y = \sin\theta\sin\lambda, \quad z = \cos\theta$$

and

$$dS = 1^2\sin\theta d\theta d\lambda$$

Whence, since $\mathbf{F}$ is given in terms of Cartesian co-ordinates, we write $\hat{\mathbf{r}} = \mathbf{i}\sin\theta\cos\lambda + \mathbf{j}\sin\theta\sin\lambda + \mathbf{k}\cos\theta$ so that $d\mathbf{S} = \hat{\mathbf{r}}\sin\theta d\theta d\lambda$. Remember that $\hat{\mathbf{r}}$, the unit position vector, is the unit normal to the unit sphere $x^2 + y^2 + z^2 = 1$ and hence the directed element of area $d\mathbf{S}$ must also be in this direction. Evaluating gives

$$\mathbf{F} \cdot d\mathbf{S} = (x^2\sin^2\theta\cos\lambda + y^2\sin^2\theta\sin\lambda + z^2\sin\theta\cos\theta)d\theta d\lambda$$

which, substituting for x, y and z on $x^2 + y^2 + z^2 = 1$ gives

$$\mathbf{F} \cdot d\mathbf{S} = (\sin^4\theta\cos^3\lambda + \sin^4\theta\sin^3\lambda + \sin\theta\cos^3\theta)d\theta d\lambda$$

So, $\int_S \mathbf{F} \cdot d\mathbf{S} = \int_0^{\pi}\int_0^{2\pi}(\sin^4\theta\cos^3\lambda + \sin^4\theta\sin^3\lambda + \sin\theta\cos^3\theta)d\theta d\lambda$, and integrating once

$$\begin{aligned}\int_S \mathbf{F} \cdot d\mathbf{S} &= \int_0^{\pi}\left[\left(\frac{3}{8}\theta - \frac{1}{4}\sin 2\theta + \frac{1}{32}\sin 4\theta\right)(\sin^3\lambda + \cos^3\lambda) - \cos\theta\cos^3\lambda\right]_0^{2\pi} d\lambda \\ &= \frac{3}{4}\pi\int_0^{\pi}(\sin^3\lambda + \cos^3\lambda)d\lambda = \frac{3}{4}\pi\left[-\cos\lambda + \frac{1}{3}\cos^3\lambda + \sin\lambda - \frac{1}{3}\sin^3\lambda\right]_0^{\pi} \\ &= \frac{3}{4}\pi \cdot 2 \cdot \frac{2}{3} = \pi\end{aligned}$$

Example 10.8 The vector fields **E** and **H** are electric and magnetic fields respectively given by

$$\mathbf{E} = \hat{\boldsymbol{\theta}}\frac{E_0}{r}\sin\theta\cos(\omega t - \beta r) \text{ and } \mathbf{H} = \hat{\boldsymbol{\lambda}}\frac{H_0}{r}\sin\theta\cos(\omega t - \beta r)$$

where E_0, H_0, ω and β are constants, and (r, θ, λ) are spherical polar co-ordinates. Find Poynting's vector $\mathbf{P} = \mathbf{E} \times \mathbf{H}$, and hence determine the instantaneous power

$$W = \iint_S \mathbf{P} \cdot d\mathbf{S}$$

over a sphere centre the origin of arbitrary radius.

Solution Since (r, θ, λ) is a right-handed system, Poynting's vector **P** is in the $\hat{\mathbf{r}}$ direction and of magnitude $\dfrac{E_0H_0}{r^2}\sin^2\theta\cos^2(\omega t - \beta r)$. If we assume that the sphere is of radius a, then the element of area of the surface of this sphere is

$$dS = a^2 \sin\theta d\theta d\lambda$$

and so the surface integral for W can be written as the following double integral in $\theta\lambda$ space

$$\iint_S \mathbf{P} \cdot d\mathbf{S} = E_0H_0\int_0^{2\pi}\int_0^{\pi}\left(\frac{\sin^2\theta}{a^2}\right)\cos^2(\omega t - \beta a)a^2 \sin\theta d\theta d\lambda$$

which simplifies to

$$\iint_S \mathbf{P} \cdot d\mathbf{S} = 2\pi E_0H_0\cos^2(\omega t - \beta a)\int_0^{\pi}\sin^3\theta d\theta$$

Since $\displaystyle\int_0^{\pi}\sin^3\theta d\theta = \int_0^{\pi}\sin\theta(1 - \cos^2\theta)d\theta = \left[-\cos\theta + \frac{1}{3}\cos^3\theta\right]_0^{\pi} = \frac{4}{3}$, we have, finally

$$W = \iint_S \mathbf{P} \cdot d\mathbf{S} = \frac{8}{3}\pi E_0H_0\cos^2(\omega t - \beta a)$$

10.3 Exercises

10.1. Evaluate the integral $\iint_S \mathbf{A} \cdot d\mathbf{S}$ for the following cases:
(a) $\mathbf{A} = y\mathbf{i} + 2x\mathbf{j} - z\mathbf{k}$ and S is the surface of the plane $2x + y = 6$ in the first octant, cut off by the plane $z = 4$.
(b) $\mathbf{A} = \mathbf{r}$ and S is the hemisphere $x^2 + y^2 + z^2 = a^2$, $z \geq 0$.
(c) $\mathbf{A} = (x + z^2)\mathbf{i} - 2x\mathbf{j} + 2yz\mathbf{k}$ and S is the surface consisting of that part of the plane $2x + y + 2z = 6$ that is in the first quadrant.

10.2. The *flux* of **F** through the surface S is defined as the integral $\iint_S \mathbf{F} \cdot d\mathbf{S}$ where S is usually a closed surface. Determine the flux of **F** through S for the following cases:
(a) $\mathbf{F} = x\mathbf{i} + y\mathbf{j}$, where S is the sphere $x^2 + y^2 + z^2 = a^2$,
(b) $\mathbf{F} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$, where S is the sphere $x^2 + y^2 + z^2 = a^2$,
(c) $\mathbf{F} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$, where S is the cube with vertices $(0, 0, 0)$, $(a, 0, 0)$, $(0, a, 0)$, $(0, 0, a)$, $(a, a, 0)$, $(0, a, a)$, $(a, 0, a)$ and (a, a, a).
(d) $\mathbf{F} = 4xz\mathbf{i} + xyz^2\mathbf{j} + 3z\mathbf{k}$, where S is the surface of the cone $z^2 = x^2 + y^2$ including its plane end $z = 4$.

10.3. Let the equation of a surface S be given by $z = f(x, y)$, and let R be the projection of this surface on to the x–y plane. Show that the area of S is given by:

$$\int_R \sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2}\, dxdy$$

10.4. Let S be that part of the paraboloid $z = 1 - x^2 - y^2$ that lies above the unit disc $x^2 + y^2 = 1$. Find the flux of **r** through S.

10.5. Evaluate by parameterisation the scalar surface integral:

$$\iint_S (x + y + z)dS$$

where S is that part of the unit sphere $x^2 + y^2 + z^2 = 1$ that lies in the positive octant.

10.6. Repeat Exercise 10.5, but this time use projection on to the quadrant $x^2 + y^2 = 1$, $x \geq 0$, $y \geq 0$. Which is the better method?

10.7. If $\mathbf{F} = \dfrac{\lambda\mathbf{r}}{|\mathbf{r}|^3}$, where λ is a constant, is the inverse square law, show that the flux of **F** through a sphere centre the origin is independent of the radius of the sphere and find its value.

10.8. If $\mathbf{F} = \dfrac{\mu\mathbf{r}}{|\mathbf{r}|^4}$, where, μ is a constant, find the flux of **F** through any sphere centre the origin. Is this result still independent of the radius of the sphere?

10.9. Calculate the value of the scalar surface integral

$$\iint_S \sqrt{x^2 + y^2}dS$$

by direct parameterisation if S is the helicoid $x = u\cos v$, $y = u\sin v$, $z = bv$, $0 \leq u \leq 1$, $0 \leq v \leq 2\pi$. (A *helicoid* has the shape of an old-fashioned Christmas streamer and resembles a spiral staircase.)

10.10. A rain gauge collector is in the form of a cone of shape $z^2 = x^2 + y^2$, having unit radius at its mouth. Treating a heavy downpour as the constant vector $\mathbf{F} = -\mathbf{k}$, determine the flux of $\mathbf{F}$ through the cone. What is the new flux if a gale blows and the rain is now at 45 degrees so that $\mathbf{F} = -\dfrac{\mathbf{i}}{\sqrt{2}} - \dfrac{\mathbf{k}}{\sqrt{2}}$.

10.11. Show that the expression $A(S)$ defined by the integral

$$A(S) = \iint_S \left[\left(\frac{\partial(x, y)}{\partial(u, v)}\right)^2 + \left(\frac{\partial(y, z)}{\partial(u, v)}\right)^2 + \left(\frac{\partial(z, x)}{\partial(u, v)}\right)^2\right]^{1/2} \mathrm{d}u\mathrm{d}v$$

is the surface area of S parameterised by $\mathbf{r} = \mathbf{r}(u, v)$. Hence determine the following surface areas:
(a) the cone $x = u\cos v$, $y = u\sin v$, $z = u$, $0 \leq v \leq 2\pi$, $0 \leq u \leq 1$,
(b) the helicoid $x = u\cos v$, $y = u\sin v$, $z = v$, $0 \leq v \leq 2\pi$, $0 \leq u \leq 1$.

10.12. If a surface is given by the equation $z = f(x, y)$ show that

$$\int_S \mathbf{F} \cdot \mathrm{d}\mathbf{S} = \int_R \left\{-\frac{\partial f}{\partial x} F_1 - \frac{\partial f}{\partial y} F_2 + F_3\right\} \mathrm{d}x\mathrm{d}y$$

where R is the projection of $z = f(x, y)$ on to the x–y plane, and $\mathbf{F} = F_1\mathbf{i} + F_2\mathbf{j} + F_3\mathbf{k}$.

Topic Guide

Gauss's and Stokes' Theorems
Green's Second Theorem
Co-ordinate Free Definitions of Div and Curl

11 Integral Theorems

11.1 Fact Sheet

If $\mathbf{F}$ is a vector-valued function with continuous partial derivatives throughout a region V, and V is surrounded by a closed surface S, then

$$\int_V \nabla \cdot \mathbf{F} \mathrm{d}V = \int_S \mathbf{F} \cdot \mathrm{d}\mathbf{S}$$

This result is known as *Gauss's Flux Theorem* or *Gauss's Divergence Theorem* (or sometimes just as the Divergence Theorem).

If $\mathbf{F}$ is a vector-valued function with continuous derivatives throughout an open surface S which itself is bordered by a simple closed surface C then

$$\int_S \nabla \times \mathbf{F} \cdot \mathrm{d}\mathbf{S} = \int_C \mathbf{F} \cdot \mathrm{d}\mathbf{s}$$

where $\mathrm{d}\mathbf{s}$ is the directed element of the curve C. This result is known as *Stokes' Theorem.*

If $\phi(x, y, z)$ and $\psi(x, y, z)$ are scalar functions of (x, y, z) with continuous second-order partial derivatives throughout a volume V with V being surrounded by a closed surface S, then

$$\int_V (\phi\nabla^2\psi - \psi\nabla^2\phi)\mathrm{d}V = \int_S (\phi\nabla\psi - \psi\nabla\phi) \cdot \mathrm{d}\mathbf{S}$$

This result is known as *Green's Second Theorem.* One of the consequences of Stokes' and Gauss's Theorems of particular delight to pure mathematicians is to be able to provide co-ordinate free definitions of divergence and curl. These are as follows:

$$\nabla \cdot \mathbf{A} = \lim_{\Delta V \to 0} \left\{\frac{1}{\Delta V}\int_S \mathbf{A} \cdot \mathrm{d}\mathbf{S}\right\}$$

and

$$(\nabla \times \mathbf{A}) \cdot \hat{\mathbf{n}} = \lim_{\Delta S \to 0} \left\{\frac{1}{\Delta S}\int_C \mathbf{A} \cdot \mathrm{d}\mathbf{s}\right\}$$

where in both definitions, the infinitesimal quantities ΔV and ΔS tend to zero in such a way that the limit exists, and in the second definition, $\hat{\mathbf{n}}$ is the unit normal to the element of area ΔS.

11.2 Worked Examples

Example 11.1 Prove Gauss's Flux Theorem: that if $\mathbf{F}$ is a vector-valued function with continuous partial derivatives throughout a region V, and V is surrounded by a closed surface S, then

$$\int_V \nabla \cdot \mathbf{F} \mathrm{d}V = \int_S \mathbf{F} \cdot \mathrm{d}\mathbf{S}$$

Solution The closed surface S is relabelled as in Figure 11.1, with S_1 denoting the lower part of it having equation $z = f_1(x, y)$ and with S_2 denoting the upper part of it having equation $z = f_2(x, y)$. This is essential to avoid multi-valued functions. The proof is along similar lines to the proof of Green's Theorem in the Plane (see Example 9.8). First of all, consider the volume integral $\int_V \frac{\partial F_3}{\partial z}\, dV$. Evaluate this by taking vertical strips as shown in Figure 11.1.

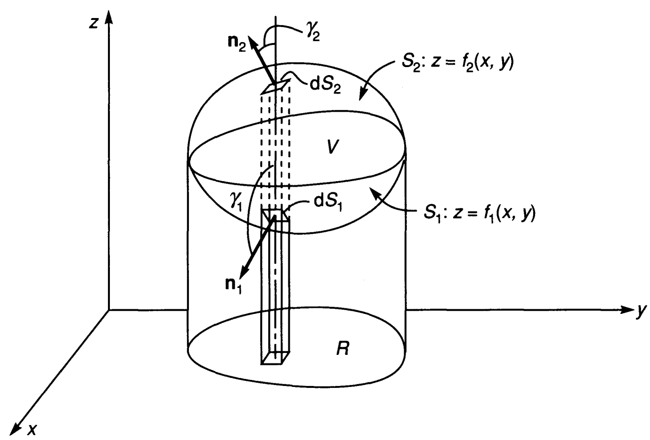


Figure 11.1 The volume V, its upper surface S_2, lower surface S_1, and vertical strip. [Adapted from M. R. Spiegel, *Vector Analysis*, unnumbered figure, page 117, Schaum, McGraw-Hill, New York, 1959]

$$\int_V \frac{\partial F_3}{\partial z}\, dV = \int_V \frac{\partial F_3}{\partial z}\, dz dx dy$$

$$= \iint_R \left\{ \int_{f_1(x,y)}^{f_2(x,y)} \frac{\partial F_3}{\partial z}\, dz \right\} dy dx$$

$$= \iint_R [F_3(x, y, f_2) - F_3(x, y, f_1)] dy dx$$

where R is shown in Figure 11.1 as the shadow cast by S on the x–y plane. Now, we do not actually want to stay on R as we would for evaluation of volume integrals, instead we want to relate the double integral over the flat surface to the integral back over the bounding surface S of V. For the upper portion of S, S_2, we have that

$$dx dy = \cos\gamma_2 dS_2 = \mathbf{k} \cdot \hat{\mathbf{n}}_2 dS_2 = \mathbf{k} \cdot d\mathbf{S}$$

since γ_2 is acute. For the lower portion of S, S_1, we have that

$$dx dy = \cos\gamma_1 dS_1 = -\mathbf{k} \cdot \hat{\mathbf{n}}_1 dS_1 = -\mathbf{k} \cdot d\mathbf{S}$$

since γ_1 is obtuse. This means that for the upper portion of S

$$\iint_R F_3(x, y, f_2) dy dx = \int_{S_2} F_3\mathbf{k} \cdot d\mathbf{S}$$

and for the lower portion

$$\iint_R F_3(x, y, f_1) dy dx = -\int_{S_1} F_3\mathbf{k} \cdot d\mathbf{S}$$

Hence

$$\int_V \frac{\partial F_3}{\partial z}\, dV = \iint_R [F_3(x, y, f_2) - F_3(x, y, f_1)] dy dx$$

$$= \int_{S_2} F_3\mathbf{k} \cdot d\mathbf{S} + \int_{S_1} F_3\mathbf{k} \cdot d\mathbf{S}$$

$$= \int_S F_3\mathbf{k} \cdot d\mathbf{S}$$

By projecting on to the other co-ordinate planes, we obtain in a similar manner:

$$\int_V \frac{\partial F_2}{\partial y}\, \mathrm{d}V = \int_S F_2\mathbf{j} \cdot \mathrm{d}\mathbf{S}$$

and
$$\int_V \frac{\partial F_1}{\partial x}\, \mathrm{d}V = \int_S F_1\mathbf{i} \cdot \mathrm{d}\mathbf{S}$$

Adding these three results gives $\int_V \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z}\, \mathrm{d}V = \int_S F_1\mathbf{i} + F_2\mathbf{j} + F_3\mathbf{k}) \cdot \mathrm{d}\mathbf{S}$ or $\int_V \boldsymbol{\nabla} \cdot \mathbf{F}\mathrm{d}V = \int_S \mathbf{F} \cdot \mathrm{d}\mathbf{S}$ which is Gauss's Flux Theorem.

Example 11.2 Prove Stokes' theorem that states: if $\mathbf{F}$ is a vector-valued function with continuous derivatives throughout an open surface S which itself is bordered by a simple closed surface C then

$$\int_S \boldsymbol{\nabla} \times \mathbf{F} \cdot \mathrm{d}\mathbf{S} = \int_C \mathbf{F} \cdot \mathrm{d}\mathbf{s}$$

where $\mathrm{d}\mathbf{s}$ is the directed element of the curve C.

Solution To prove this result we use a similar approach to the previous example. However, later in the proof, things do get a little technical so you need to pay attention; it is not just a case of turning the handle! First we split $\mathbf{F}$ into components $\mathbf{F} = F_1\mathbf{i} + F_2\mathbf{j} + F_3\mathbf{k}$ and consider the curl of one of its components, say $\boldsymbol{\nabla} \times (F_1\mathbf{i})$. Now, $\boldsymbol{\nabla} \times (F_1\mathbf{i}) = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_1 & 0 & 0 \end{vmatrix} = \frac{\partial F_1}{\partial z}\mathbf{j} - \frac{\partial F_1}{\partial y}\mathbf{k}$. Also, suppose that $z = f(x, y)$ is the equation of the open surface S so that $\mathbf{r}$, the position vector of any point on S, is given by $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k} = x\mathbf{i} + y\mathbf{j} + f(x, y)\mathbf{k}$. It turns out that we shall need the derivative $\frac{\partial \mathbf{r}}{\partial y} = \mathbf{j} + \frac{\partial f}{\partial y}\mathbf{k}$, which incidentally is perpendicular to $\hat{\mathbf{n}}$, the normal to the surface S since $\frac{\partial \mathbf{r}}{\partial y}$ itself is tangent to S. Hence we have the result $\hat{\mathbf{n}} \cdot \mathbf{j} + \frac{\partial f}{\partial y}\hat{\mathbf{n}} \cdot \mathbf{k} = 0$ which will prove useful. Now we have already shown that

$$\boldsymbol{\nabla} \times (F_1\mathbf{i}) = \frac{\partial F_1}{\partial z}\,\mathbf{j} - \frac{\partial F_1}{\partial y}\,\mathbf{k}$$

and taking the scalar product of this with $\hat{\mathbf{n}}$ gives

$$\begin{aligned} \boldsymbol{\nabla} \times (F_1\mathbf{i}) \cdot \hat{\mathbf{n}} &= \frac{\partial F_1}{\partial z}\,\mathbf{j} \cdot \hat{\mathbf{n}} - \frac{\partial F_1}{\partial y}\,\mathbf{k} \cdot \hat{\mathbf{n}} \\ &= \mathbf{k} \cdot \hat{\mathbf{n}}\left\{-\frac{\partial f}{\partial y}\frac{\partial F_1}{\partial z} - \frac{\partial F_1}{\partial y}\right\} \end{aligned}$$

using $\hat{\mathbf{n}} \cdot \mathbf{j} + \frac{\partial f}{\partial y}\,\hat{\mathbf{n}} \cdot \mathbf{k} = 0$ to substitute $\mathbf{j} \cdot \hat{\mathbf{n}} = -\frac{\partial f}{\partial y}\,\mathbf{k} \cdot \hat{\mathbf{n}}$. All this is straightforward if a little obscure. The next step demands understanding of the properties of partial derivatives as used in Chapter 2. On S

$$F_1(x, y, z) = F_1(x, y, f(x, y)) = G(x, y) \text{ (say)}$$

hence using the chain rule for partial differentiation

$$\frac{\partial G}{\partial y} = \frac{\partial F_1}{\partial y} + \frac{\partial F_1}{\partial f}\frac{\partial f}{\partial y} = \frac{\partial F_1}{\partial y} + \frac{\partial F_1}{\partial z}\frac{\partial f}{\partial y} \quad \text{since } \frac{\partial F_1}{\partial f} = \frac{\partial F_1}{\partial z} \text{ on } S.$$

So $\quad \nabla \times (F_1\mathbf{i}) \cdot \hat{\mathbf{n}} = -\dfrac{\partial G}{\partial y}\mathbf{k} \cdot \hat{\mathbf{n}}$

and hence $\quad \nabla \times (F_1\mathbf{i}) \cdot \hat{\mathbf{n}}\mathrm{d}S = -\dfrac{\partial G}{\partial y}\mathbf{k} \cdot \hat{\mathbf{n}}\mathrm{d}S = -\dfrac{\partial G}{\partial y}\mathrm{d}x\mathrm{d}y$

using the projection of S on to the x–y plane (see Figure 11.2). Hence at last we can use the projection form of evaluating a surface integral (see Example 10.1) to write

$$\int_S \nabla \times F_1\mathbf{i} \cdot \mathrm{d}\mathbf{S} = \int_R -\frac{\partial G}{\partial y}\,\mathrm{d}x\mathrm{d}y$$

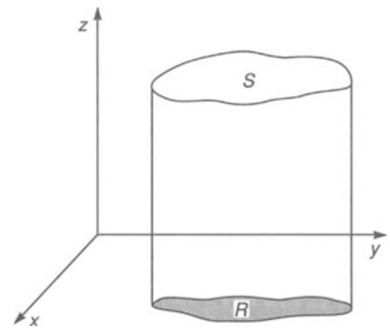


Figure 11.2 The projection of S on to the x–y plane.

Using Green's Theorem in the Plane, we have

$$\int_R -\frac{\partial G}{\partial y}\,\mathrm{d}x\mathrm{d}y = \int_\Gamma G\mathrm{d}x$$

where Γ is the boundary of R. However, at each point of Γ, $G(x, y)$ has the same value as $F_1(x, y)$ has on the corresponding point of C (the bounding curve of S).

Hence $\int_S \nabla \times F_1\mathbf{i} \cdot \mathrm{d}\mathbf{S} = \int_\Gamma G\mathrm{d}x = \int_C F_1\mathrm{d}x$

Similarly we can deduce $\int_S \nabla \times F_2\mathbf{j} \cdot \mathrm{d}\mathbf{S} = \int_C F_2\mathrm{d}y$ and $\int_S \nabla \times F_3\mathbf{k} \cdot \mathrm{d}\mathbf{S} = \int_C F_3\mathrm{d}z$

Adding these results gives $\int_S \nabla \times \mathbf{F} \cdot \mathrm{d}\mathbf{S} = \int_C \mathbf{F} \cdot \mathrm{d}\mathbf{S}$ which is Stokes' Theorem. (The use of x, y and z for both the elemental lengths of Γ and C is allowable licence for an applied mathematician and, hopefully, minimises confusion). See Exercise 11.6 for an alternative proof.

Example 11.3 Use Stokes' Theorem to show that a necessary and sufficient condition for $\oint_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} = 0$ for *every* closed curve C in a region R is that $\nabla \times \mathbf{F} = \mathbf{0}$ throughout R. Deduce the value of $\int_{P_1P_2} \mathbf{F} \cdot \mathrm{d}\mathbf{r}$ if $\nabla \times \mathbf{F} = \mathbf{0}$ everywhere inside R. The integral $\int_{P_1P_2}$ denotes the integral along any path connecting P_1 and P_2, where P_1 and P_2 are arbitrary points of R.

Solution This problem provides an example of proving a condition both necessary and sufficient. As is usually the case, one is easier than the other; in this instance it is sufficiency that is easier. The result itself is fundamental to potential theory and thence to any physical system that can be described in terms of a potential.

To prove sufficiency, assume that $\nabla \times \mathbf{F} = \mathbf{0}$, then we have by Stokes' Theorem

$$\oint_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \int_R \nabla \times \mathbf{F} \cdot \mathrm{d}\mathbf{S} = 0$$

which shows that $\oint_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} = 0$ for *every* closed curve C in a region R. That was quick! To prove necessity is a little slower.

Now, suppose that $\oint_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} = 0$ for *every* closed curve C in a region R, and let us assume that there is a point inside R for which $\nabla \times \mathbf{F} \neq \mathbf{0}$. This means that there is a non-zero constant α such that $\nabla \times \mathbf{F} = \alpha\hat{\mathbf{p}}$. We can design a surface S that passes through this point and is such that its normal is in the same direction as $\nabla \times \mathbf{F}$.

Whence we have the result $\oint_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \int_S \nabla \times \mathbf{F} \,.\, \mathrm{d}\mathbf{S} = \int_S \alpha\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}\mathrm{d}S = \alpha \int_S \mathrm{d}S$ and this result is not zero since α is non-zero. This contradicts the initial supposition $\oint_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} = 0$ and we have a contradiction. Thus the assumption that there is a point inside R for which $\nabla \times \mathbf{F} \neq \mathbf{0}$ must be false and we are forced to conclude that $\nabla \times \mathbf{F} = \mathbf{0}$ at every point of R, which proves necessity. It is typical that the harder way round to prove a result has all the elegance of being pulled through a hedge backwards. It is nevertheless no less valid.

Example 11.4 Prove that if $\phi(x, y, z)$ and $\psi(x, y, z)$ are scalar functions of (x, y, z) with continuous second-order partial derivatives throughout a volume V with V being surrounded by a closed surface S, then

$$\int_V (\phi\nabla^2\psi - \psi\nabla^2\phi)\mathrm{d}V = \int_S (\phi\nabla\psi - \psi\nabla\phi) \cdot \mathrm{d}\mathbf{S}$$

(Green's Second Theorem).

Solution This theorem is proved reasonably easily by putting $\mathbf{F} = \phi\nabla\psi$ in the Divergence Theorem. This gives

$$\int_V \nabla \cdot (\phi\nabla\psi)\mathrm{d}V = \int_S \phi\nabla\psi \cdot \mathrm{d}\mathbf{S}$$

However, we have the identity $\nabla \cdot (\phi\nabla\psi) = \phi\nabla^2\psi + \nabla\phi \cdot \nabla\psi$ so, for example, $\nabla \cdot (\phi\nabla\psi) - \nabla \cdot (\psi\nabla\phi) = \phi\nabla^2\psi - \psi\nabla^2\phi$ or $\nabla \cdot (\phi\nabla\psi - \psi\nabla\phi) = \phi\nabla^2\psi - \psi\nabla^2\phi$. The Divergence Theorem applied to the vector field $(\phi\nabla\psi - \psi\nabla\phi)$ gives the desired result as follows:

$$\int_V \nabla \cdot (\phi\nabla\psi - \psi\nabla\phi)\mathrm{d}V = \int_V \phi\nabla^2\psi - \psi\nabla^2\phi\mathrm{d}V = \int_S (\phi\nabla\psi - \psi\nabla\phi) \cdot \mathrm{d}\mathbf{S}$$

Example 11.5 If $\phi = \phi(x, y, z)$ is a scalar function of position with continuous second-order partial derivatives inside and on a domain D in which it also satisfies the equation $\nabla^2\phi = 0$, show that $\phi = \phi(x, y, z)$ is unique apart from perhaps a multiplicative constant.

Solution This result is fundamental in potential theory and leads to the ability to solve problems in inviscid flow, electromagnetism and linear elasticity as well as in other fields with an assurance of uniqueness. Suppose there are two fields ϕ and ψ that satisfy Laplace's equation, that is $\nabla^2\phi = 0$ and $\nabla^2\psi = 0$. The left-hand side of Green's Second Theorem is thus zero, hence

$$0 = \int_V \phi\nabla^2\psi - \psi\nabla^2\phi\mathrm{d}V = \int_S (\phi\nabla\psi - \psi\nabla\phi) \cdot \mathrm{d}\mathbf{S}$$

from which, since $\nabla\phi \cdot \mathrm{d}\mathbf{S} = \dfrac{\partial\phi}{\partial n}\mathrm{d}S$ and $\nabla\psi \cdot \mathrm{d}\mathbf{S} = \dfrac{\partial\psi}{\partial n}\mathrm{d}S$

$$\int_S \left(\phi\frac{\partial\psi}{\partial n} - \psi\frac{\partial\phi}{\partial n}\right)\mathrm{d}S = 0$$

Now, since S is arbitrary, we must have

$$\phi\frac{\partial\psi}{\partial n} - \psi\frac{\partial\phi}{\partial n} = 0$$

or, rearranging

$$\frac{1}{\psi}\frac{\partial\psi}{\partial n} = \frac{1}{\phi}\frac{\partial\phi}{\partial n}$$

so that $$\frac{\partial}{\partial n}(\ln\psi) = \frac{\partial}{\partial n}(\ln\phi)$$

from which, upon integrating $$\ln\psi = \ln\phi + \text{const.}$$

or $$\psi = c\phi$$

which implies that ϕ and ψ are equal, apart from perhaps a multiplicative constant.

Example 11.6 If $\mathbf{F}$ is a vector field with continuous derivatives throughout a volume V which is bounded by a closed surface S show that $\int_V \nabla \times \mathbf{F} dV = \mathbf{0}$ if $\mathbf{F}$ is everywhere normal to the surface S.

Solution Whenever volumes and closed bounding surfaces are involved, the mind should turn to Gauss's Divergence Theorem $\int_V \nabla \cdot \mathbf{A} dV = \int_S \mathbf{A} \cdot d\mathbf{S}$. However, the question involves the curl therefore we try to derive a like expression, so let $\mathbf{A} = \mathbf{F} \times \mathbf{C}$ and consider $\nabla \cdot (\mathbf{F} \times \mathbf{C})$ where $\mathbf{C}$ is a constant vector. Using the identity

$$\nabla \cdot (\mathbf{F} \times \mathbf{C}) = \mathbf{C} \cdot \nabla \times \mathbf{F} - \mathbf{F} \cdot \nabla \times \mathbf{C}$$

we have $\qquad \nabla \cdot (\mathbf{F} \times \mathbf{C}) = \mathbf{C} \cdot \nabla \times \mathbf{F}$ since $\mathbf{C}$ is a constant vector.

Gauss's Divergence Theorem applied to the vector $\mathbf{F} \times \mathbf{C}$ thus gives

$$\int_V \nabla \cdot (\mathbf{F} \times \mathbf{C}) dV = \int_S \mathbf{F} \times \mathbf{C} \cdot d\mathbf{S}$$

The right-hand side is a scalar triple product, albeit one involving the infinitesimal quantity $d\mathbf{S}$. Thus we can use $\mathbf{F} \times \mathbf{C} \cdot d\mathbf{S} = d\mathbf{S} \cdot \mathbf{F} \times \mathbf{C} = \mathbf{C} \cdot d\mathbf{S} \times \mathbf{F}$, and hence $\int_V \nabla \cdot (\mathbf{F} \times \mathbf{C}) dV = \int_V \mathbf{C} \cdot \nabla \times \mathbf{F} dV = \mathbf{C} \cdot \int_V \nabla \times \mathbf{F} dV = \int_S \mathbf{F} \times \mathbf{C} \cdot d\mathbf{S}$ (the last equality is from Gauss's Flux Theorem applied to the vector $\mathbf{F} \times \mathbf{C}$).
This left-hand side is the same as $\int_S \mathbf{C} \cdot d\mathbf{S} \times \mathbf{F} = \mathbf{C} \cdot \int_S d\mathbf{S} \times \mathbf{F}$. Hence Gauss's Theorem implies $\mathbf{C} \cdot \left[\int_V \nabla \times \mathbf{F} dV - \int_S d\mathbf{S} \times \mathbf{F}\right] = 0$. The constant vector $\mathbf{C}$ is an arbitrary choice, therefore we cannot deduce that it is always perpendicular to the square bracket and instead it must be true that

$$\begin{aligned} \int_V \nabla \times \mathbf{F} dV &= \int_S d\mathbf{S} \times \mathbf{F} \\ &= \int_S \hat{\mathbf{n}} \times \mathbf{F} dS \\ &= \mathbf{0} \end{aligned}$$

since $\mathbf{F}$ is parallel to $\hat{\mathbf{n}}$ (we are given that $\mathbf{F}$ is perpendicular to S).

Example 11.7 Show that the area enclosed by a curve C is given by $\frac{1}{2} \oint_C (x dy - y dx)$ and use this formula to calculate the area of an ellipse.

Solution The area of a plane surface is $\int_S dS = \int_S dx dy$ in Cartesian co-ordinates. Green's Theorem in the Plane is $\int_C P dx + Q dy = \int_S \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx dy$. Putting $Q = x$ and $P = -y$ gives the right-hand side of Green's Theorem as $2\int_S dS = 2\int_S dx dy$. Hence we have via Green's Theorem $\int_S dx dy = \frac{1}{2}\int_C (x dy - y dx)$ as the surface area enclosed by C. The parametric representation of the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ is $x = a\cos\theta$, $y = b\sin\theta$ where $0 \leq \theta < 2\pi$. On C, $dx = -a\sin\theta d\theta$, $dy = b\cos\theta d\theta$ and so we have $x dy - y dx = [ab\cos^2\theta - b\sin\theta(-a\sin\theta)]d\theta = ab(\cos^2\theta + \sin^2\theta)d\theta = ab d\theta$. Thus the area of the ellipse is $\frac{1}{2}\int_C (x dy - y dx) = \frac{1}{2}\int_0^{2\pi} ab d\theta = \pi ab$.

Example 11.8 Establish, using the previous example, the two alternative expressions for the area enclosed by a simple closed curve C: $\frac{1}{2}\int_C x^2 d\left(\frac{y}{x}\right)$, $\frac{1}{2}\int_C R^2 d\theta$. Hence calculate the area inside the Folium of Descartes $x^3 + y^3 = 3axy$ $(a > 0)$ and the area enclosed by the cardioid $R = a(1 + \cos\theta)$; (R, θ) are plane polar co-ordinates.

Solution The expression derived in the last example was $\int_S \mathrm{d}x\mathrm{d}y = \frac{1}{2}\int_C (x\mathrm{d}y - y\mathrm{d}x)$. Note that $\mathrm{d}\left(\frac{y}{x}\right) = \frac{x\mathrm{d}y - y\mathrm{d}x}{x^2}$, hence $x^2\mathrm{d}\left(\frac{y}{x}\right) = x\mathrm{d}y - y\mathrm{d}x$ and so the area enclosed by C is $\frac{1}{2}\int_C x^2\mathrm{d}\left(\frac{y}{x}\right)$ as required. Putting $y = R\sin\theta$, $x = R\cos\theta$ gives $\frac{y}{x} = \tan\theta$ so that $\mathrm{d}\left(\frac{y}{x}\right) = \sec^2\theta\mathrm{d}\theta$ and so $x^2\mathrm{d}\left(\frac{y}{x}\right) = R^2\cos^2\theta\sec^2\theta\mathrm{d}\theta = R^2\mathrm{d}\theta$. Thus the area is also given by $\frac{1}{2}\int_C R^2\mathrm{d}\theta$ as required.

In order to calculate the area inside the Folium of Descartes we need to parameterise the curve. The most straightforward way of doing this is to put $y = xt$. The Folium has equation $x^3 + y^3 = 3axy$, whence $x^3 + x^3t^3 = 3ax^2t$ and so $x = \frac{3at}{1 + t^3}$ and $y = \frac{3at^2}{1 + t^3}$, which with $\frac{y}{x} = t$ gives $x^2\mathrm{d}\left(\frac{y}{x}\right) = \frac{9a^2t^2}{(1 + t^3)^2}\mathrm{d}t$. The shape of the curve is shown in Figure 11.3.

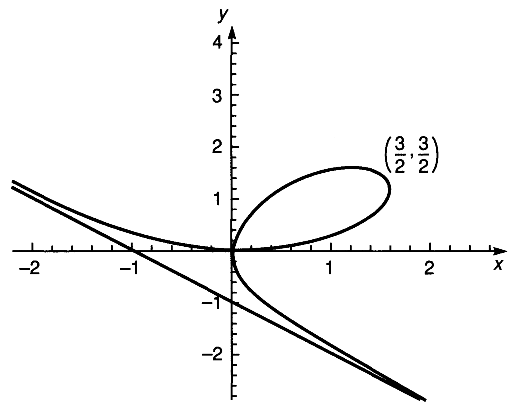


Figure 11.3 The Folium of Descartes $x^3 + y^3 = 3axy$ (with $a = 1$).

Utilising $x = \frac{3at}{1 + t^3}$ and $y = \frac{3at^2}{1 + t^3}$ we note that if $t = 0$, $x = 0$, $y = 0$ and as $t \to \infty$ once again $x \to 0$, $y \to 0$. The point 'half-way round' the loop shown in Figure 11.3 is $\left(\frac{3}{2}a, \frac{3}{2}a\right)$ and corresponds to $t = 1$. Therefore, as t varies from 0 to ∞ the loop is circuited once, thus $\int_C x^2\mathrm{d}\left(\frac{y}{x}\right) = \int_0^\infty\left(\frac{3at}{1 + t^3}\right)^2\mathrm{d}t = 9a^2\int_0^\infty \frac{t^2}{(1 + t^3)^2}\mathrm{d}t$. Evaluating this by substituting $u = 1 + t^3$, $t^2\mathrm{d}t = \frac{1}{3}\mathrm{d}u$ we we proceed as follows:

$$\text{Area of the Folium} = \tfrac{1}{2}9a^2\,\tfrac{1}{3}\int_1^\infty \frac{\mathrm{d}u}{u^2} = \frac{3a^2}{2}\left[-\frac{1}{u}\right]_1^\infty = \frac{3a^2}{2}$$

In order to find the area enclosed by the cardioid $R = a(1 + \cos\theta)$ (see Figure 11.4) we use the formula in polar co-ordinates, that is $\frac{1}{2}\int_C R^2\mathrm{d}\theta$ and there is no need for parameterisation with $0 \le \theta \le 2\pi$. Hence the area is given by

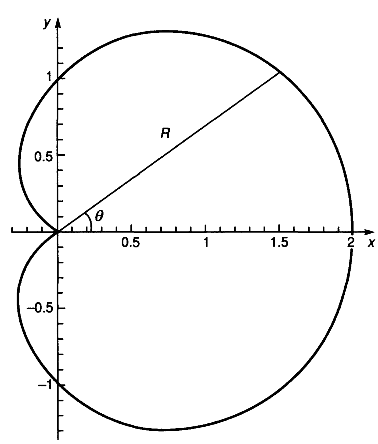


Figure 11.4 The cardioid $R = a\,(1 + \cos\theta)$[with $a = 1$].

$$\tfrac{1}{2}\int_C R^2 d\theta = \tfrac{1}{2}\int_0^{2\pi} a^2(1 + \cos\theta)^2 d\theta$$

$$= \tfrac{1}{2}\int_0^{2\pi} a^2(1 + 2\cos\theta + \tfrac{1}{2}(1 + \cos 2\theta))\, d\theta$$

$$= \tfrac{1}{2}a^2[\theta + 2\sin\theta + \tfrac{1}{2}\theta + \tfrac{1}{4}\sin 2\theta]_0^{2\pi} = \frac{3}{2}\, a^2\pi$$

coincidentally π times the area of the Folium of Descartes.

Example 11.9 If $\phi(x, y)$ is a plane harmonic function in the region S: $x^2 + y^2 \le r^2$, show that $\int_S (\nabla\phi)^2 dS = \tfrac{1}{2}r \dfrac{d}{dr}\left(\int_0^{2\pi} \phi^2 d\theta\right)$ and verify this formula for $\phi = x$.

Solution This is a theoretical problem that demands familiarity with vector identities, as well as the application of integral theorems. First of all, recall the vector identity

$$\nabla \cdot (\alpha \mathbf{F}) = \mathbf{F} \cdot \nabla\alpha + \nabla \cdot \mathbf{F}$$

and put $\alpha = \phi$, $\mathbf{F} = \nabla\phi$ to give

$$\nabla \cdot (\phi\nabla\phi) = (\nabla\phi)^2 + \phi\nabla^2\phi = (\nabla\phi)^2$$

since ϕ is harmonic. (ϕ is harmonic if $\nabla^2\phi = 0$, that is it obeys Laplace's equation.) We will return to this equation. Gauss's Flux Theorem gives

$$\int_V \nabla \cdot (\phi\nabla\phi) dV = \int_{S_1} \phi\nabla\phi \cdot d\mathbf{S}_1 = \int_{S_1} \phi \frac{\partial\phi}{\partial n}\, dS_1$$

where S_1 is the (closed) surface that surrounds the volume V with normal co-ordinate n. Hence, using the equation just derived we have

$$\int_V (\nabla\phi)^2 dV = \int_{S_1} \phi \frac{\partial\phi}{\partial n}\, dS_1$$

Consider V as a cylinder, of unit length, with axis $\mathbf{k}$ (the z-axis) and radius r. The question defines S as the disc $x^2 + y^2 \le r^2$, hence $V = 1 \times S$, $dV = 1 \cdot dS = r d\theta$ and $n = r$ so that

$$\int_S (\nabla\phi)^2 dS = \int_0^{2\pi} \phi \frac{\partial\phi}{\partial r}\, r d\theta$$

since θ varies from 0 to 2π in order to circumnavigate the cylinder (see Figure 11.5).

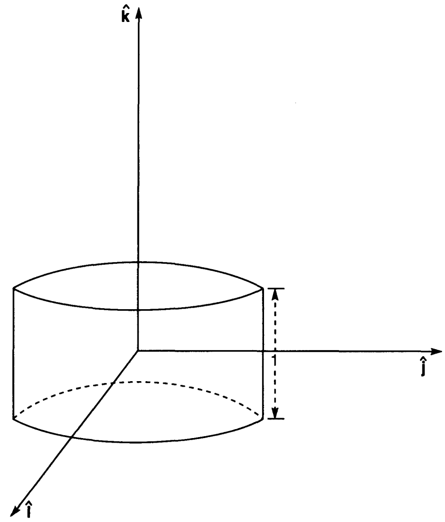


Figure 11.5 The cylinder of unit length, axis $\hat{\mathbf{k}}$.

As the integration on the right-hand side involves θ only, we can write

$$\int_0^{2\pi} \phi \frac{\partial \phi}{\partial r} r\mathrm{d}\theta = r\int_0^{2\pi} \phi \frac{\partial \phi}{\partial r} \mathrm{d}\theta = \frac{r}{2}\int_0^{2\pi} \frac{\partial}{\partial r} (\phi)^2 \, \mathrm{d}\theta = \frac{1}{2}\, r \frac{\mathrm{d}}{\mathrm{d}r}\left(\int_0^{2\pi} \phi^2 \mathrm{d}\theta\right)$$

where we have used differentiation under the integral sign (see Example 2.12). This proves the first part.

With $\phi = x$, $\boldsymbol{\nabla}\phi = \mathbf{i}$ and $(\boldsymbol{\nabla}\phi)^2 = (\boldsymbol{\nabla}\phi) \cdot (\boldsymbol{\nabla}\phi) = 1$. The left-hand side is thus $\int_S \mathrm{d}S$ the area of the circle S, that is πr^2. The right-hand side is a little more involved:

$$\begin{aligned}\frac{1}{2}\, r \frac{\mathrm{d}}{\mathrm{d}r}\left(\int_0^{2\pi} r^2\cos^2\theta \mathrm{d}\theta\right) &= \frac{1}{2}\, r \frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\int_0^{2\pi} \cos^2\theta \mathrm{d}\theta\right)\\ &= r^2 \int_0^{2\pi} \cos^2\theta\mathrm{d}\theta\\ &= r^2 \left[\tfrac{1}{2}\,\theta + \tfrac{1}{4}\sin 2\theta\right]_0^{2\pi}\\ &= \pi r^2\end{aligned}$$

hence verifying the established result.

Example 11.10 If $\phi(x, y, z)$ is a scalar function of position and $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$, show that $\int_C \phi\mathrm{d}\mathbf{r} = \int_S (\mathrm{d}\mathbf{S} \times \boldsymbol{\nabla}\phi)$ and $\int_C \mathbf{r} \times \mathrm{d}\mathbf{r} = 2\int_S \mathrm{d}\mathbf{S}$ where the open surface S in three-dimensional space is bounded by the smooth closed curve C.

Solution The place to start a problem linking an open surface with its outer contour is with Stokes' Theorem: $\int_S \boldsymbol{\nabla} \times \mathbf{F} \cdot \mathrm{d}\mathbf{S} = \int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r}$ where $\mathbf{F}$ is a suitably well behaved vector field.

Also, recall the vector identity $\boldsymbol{\nabla} \times (\phi\mathbf{a}) = \phi\boldsymbol{\nabla} \times \mathbf{a} + \boldsymbol{\nabla}\phi \times \mathbf{a} = \boldsymbol{\nabla}\phi \times \mathbf{a}$ if $\mathbf{a}$ is a constant. Applying Stokes' Theorem we have

$$\int_S \boldsymbol{\nabla}\phi \times \mathbf{a} \cdot \mathrm{d}\mathbf{S} = \int_S \boldsymbol{\nabla} \times (\phi\mathbf{a}) \cdot \mathrm{d}\mathbf{S} = \int_C \phi\mathbf{a} \cdot \mathrm{d}\mathbf{r}$$

Now, the left-hand side contains a scalar triple product so we have that

$$\begin{aligned}\boldsymbol{\nabla}\phi \times \mathbf{a} \cdot \mathrm{d}\mathbf{S} &= \mathbf{a} \times \mathrm{d}\mathbf{S} \cdot \boldsymbol{\nabla}\phi\\ &= \mathbf{a} \cdot \mathrm{d}\mathbf{S} \times \boldsymbol{\nabla}\phi\end{aligned}$$

whence

$$\int_C \phi\mathbf{a} \cdot \mathrm{d}\mathbf{r} = \int_S \mathbf{a} \cdot \mathrm{d}\mathbf{S} \times \boldsymbol{\nabla}\phi$$

Since $\mathbf{a}$ is a constant vector, this implies

$$\mathbf{a} \cdot \int_C \phi\mathrm{d}\mathbf{r} = \mathbf{a} \cdot \int_S (\mathrm{d}\mathbf{S} \times \boldsymbol{\nabla}\phi)$$

or

$$\mathbf{a} \cdot \left(\int_C \phi\mathrm{d}\mathbf{r} - \int_S (\mathrm{d}\mathbf{S} \times \boldsymbol{\nabla}\phi)\right) = 0$$

Since $\mathbf{a}$ is an arbitrary vector it is not always perpendicular to the term in parentheses; the term in parentheses must be zero, that is

$$\int_C \phi\mathrm{d}\mathbf{r} = \int_S (\mathrm{d}\mathbf{S} \times \boldsymbol{\nabla}\phi)$$

as required.

The second part is solved by considering the scalar product of the left-hand side $\int_C \mathbf{r} \times \mathrm{d}\mathbf{r}$ with an arbitrary constant vector $\mathbf{a}$. That is

$\mathbf{a} \cdot \int_C \mathbf{r} \times \mathrm{d}\mathbf{r} = \int_C \mathbf{a} \cdot \mathbf{r} \times \mathrm{d}\mathbf{r} = \int_C \mathbf{a} \times \mathbf{r} \cdot \mathrm{d}\mathbf{r} = \int_S \boldsymbol{\nabla} \times (\mathbf{a} \times \mathbf{r}) \cdot \mathrm{d}\mathbf{S}$ using Stokes' Theorem.

Now recall the identity $\boldsymbol{\nabla} \times (\mathbf{F} \times \mathbf{G}) = (\mathbf{G} \cdot \boldsymbol{\nabla})\mathbf{F} - \mathbf{G}(\boldsymbol{\nabla} \cdot \mathbf{F}) - (\mathbf{F} \cdot \boldsymbol{\nabla})\mathbf{G} + \mathbf{F}(\boldsymbol{\nabla} \cdot \mathbf{G})$ from Chapter 7 and put $\mathbf{F} = \mathbf{a}$, and $\mathbf{G} = \mathbf{r}$ where $\mathbf{a}$ is a constant so the first two terms on the right-hand side are zero. Thus we have $\boldsymbol{\nabla} \times (\mathbf{a} \times \mathbf{r}) = -(\mathbf{a} \cdot \boldsymbol{\nabla})\mathbf{r} + \mathbf{a}(\boldsymbol{\nabla} \cdot \mathbf{r})$. This simplifies further because

$\nabla \cdot \mathbf{r} = \frac{\partial x}{\partial x} + \frac{\partial y}{\partial y} + \frac{\partial z}{\partial z} = 1 + 1 + 1 = 3$ and,

writing $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$ we have

$(\mathbf{a} \cdot \nabla)\mathbf{r} = \left(a_1 \frac{\partial}{\partial x} + a_2 \frac{\partial}{\partial y} + a_3 \frac{\partial}{\partial z}\right)\mathbf{r} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k} = \mathbf{a}$, so that

$\nabla \times (\mathbf{a} \times \mathbf{r}) = -\mathbf{a} + \mathbf{a} \cdot 3 = 2\mathbf{a}$. We have thus shown that

$\mathbf{a} \cdot \int_C \mathbf{r} \times \mathbf{dr} = \int_S \nabla \times (\mathbf{a} \times \mathbf{r}) \cdot \mathbf{dS} = \int_S 2\mathbf{a} \cdot \mathbf{dS} = \mathbf{a} \cdot \int_S 2\mathbf{dS}$ and since the constant vector $\mathbf{a}$ is arbitrary, we must have that $\int_C \mathbf{r} \times \mathbf{dr} = \int_S 2\mathbf{dS}$ as required.

Example 11.11 If S is a closed surface, show that

$$\int_S \frac{\mathbf{n} \cdot \mathbf{r}}{r^3}\, dS = \begin{cases} 0 \text{ if the origin is outside } S \\ 2\pi \text{ if the origin is on } S \\ 4\pi \text{ if the origin is inside } S \end{cases}$$

Solution This result is sometimes known as Gauss's Theorem. (One of the problems if you are as prolific and clever as Karl Wilhelm Gauss is that you get many results named after you!) First of all we use Gauss's Divergence Theorem with $\mathbf{F} = \frac{\mathbf{r}}{r^3}$ to establish that

$$\int_S \frac{\mathbf{n} \cdot \mathbf{r}}{r^3}\, dS = \int_V \nabla \cdot \left(\frac{\mathbf{r}}{r^3}\right) dV$$

Now, $\nabla \cdot \left(\frac{\mathbf{r}}{r^3}\right) = \nabla\left(\frac{1}{r^3}\right) \cdot \mathbf{r} + \frac{1}{r^3}\nabla \cdot \mathbf{r}$ (the 'product rule' for divergence), and using $\nabla\left(\frac{1}{r^3}\right) = -\frac{3}{r^5}\mathbf{r}$ together with $\nabla \cdot \mathbf{r} = 3$ gives $\nabla \cdot \left(\frac{\mathbf{r}}{r^3}\right) = -\frac{3}{r^3} + \frac{1}{r^3}3 = 0$ provided $r \neq 0$ anywhere inside V. This is so if the origin is outside S which establishes the first result. That is $\int_S \frac{\mathbf{n} \cdot \mathbf{r}}{r^3}\, dS = 0$, origin outside S. The next two parts of this problem involve surrounding the origin with a small surface. This kind of procedure will be familiar to those who have studied complex variables up to and including Cauchy's Theorem, but will be new to most. The idea is simple enough. Since the origin is the only troublesome point, exclude it by surrounding it by a small sphere of radius ε. The volume V is thus made up of the small sphere of radius ε, call it V_ε, and the remainder τ say. Let the surfaces of the volumes be denoted by S_ε and σ respectively. Hence $S = \sigma + S_\varepsilon$ and

$$\int_S \frac{\mathbf{n} \cdot \mathbf{r}}{r^3}\, dS = \int_{S_\varepsilon} \frac{\mathbf{n} \cdot \mathbf{r}}{r^3}\, dS_\varepsilon + \int_\sigma \frac{\mathbf{n} \cdot \mathbf{r}}{r^3}\, d\sigma$$

the last integral being zero for both the troublesome cases simply because σ excludes the origin by construction. Let us deal with the integral over S_ε, first for the case when the origin is inside V. On the sphere S_ε we have $\mathbf{n} = \hat{\mathbf{r}}$ and $r = \varepsilon$. Also there is no dependence on any other variable than r hence

$$\int_{S_\varepsilon} \frac{\mathbf{n} \cdot \mathbf{r}}{r^3}\, dS_\varepsilon = \left.\frac{\hat{\mathbf{r}} \cdot \mathbf{r}}{r^3}\right|_{r=\varepsilon} \cdot 4\pi\varepsilon^2 = 4\pi$$

since the integrand is constant over the surface of the small sphere S_ε. Similarly, the small sphere surrounds the origin even if it is on the surface of the volume V. However this time it is 'surrounded' only by a *hemisphere* inside the original volume V, the other half being outside V. The calculation of integral over S_ε thus proceeds similarly as follows:

$$\int_{S'_\varepsilon} \frac{\mathbf{n} \cdot \mathbf{r}}{r^3}\, dS'_\varepsilon = \left.\frac{\hat{\mathbf{r}} \cdot \mathbf{r}}{r^3}\right|_{r=\varepsilon} \cdot 2\pi\varepsilon^2 = 2\pi$$

where the dash denotes that we are only integrating over that half of the sphere that lies inside the volume V. We are now at liberty to let the small sphere tend to zero, and since all results are independent of the radius of the sphere ε we have proved our result.

Example 11.12 Maxwell's equations of electromagnetism in free space with no current take the form:

$$\nabla \times \mathbf{H} = \frac{\partial \mathbf{E}}{\partial t}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{H}}{\partial t}$$

$$\nabla \cdot \mathbf{E} = \rho$$

$$\nabla \cdot \mathbf{H} = 0$$

where $\mathbf{E}$ is the electric field, $\mathbf{H}$ is the magnetic field, and ρ is the electric charge density. Use Stokes' Theorem and Gauss's Theorem to interpret these equations in terms of physical laws.

Solution The first equation, written in integral form is

$$\int_S \nabla \times \mathbf{H} \cdot \mathrm{d}\mathbf{S} = \int_S \frac{\partial \mathbf{E}}{\partial t} \cdot \mathrm{d}\mathbf{S}$$

or, using Stokes' Theorem

$$\oint_C \mathbf{H} \cdot \mathrm{d}\mathbf{r} = \frac{\partial}{\partial t}\left(\int_S \mathbf{E} \cdot \mathrm{d}\mathbf{S}\right)$$

where C is the bounding curve of the surface S. This is a statement of Ampère's circuit law which states that if there is a magnetic field of strength $\mathbf{H}$ present in a curve C that encloses a surface S, then this must have been induced by a flux of an electric field $\mathbf{E}$. Quite often, Ampère's law is stated in terms of the magnetic field induced by a current I that flows through a wire, in which case:

$$\oint_C \mathbf{H} \cdot \mathrm{d}\mathbf{r} = I$$

which is equivalent. The second equation written in integral form is

$$\int_S \nabla \times \mathbf{E} \cdot \mathrm{d}\mathbf{S} = -\frac{\partial}{\partial t}\left(\int_S \mathbf{H} \cdot \mathrm{d}\mathbf{S}\right)$$

or, using Stokes' Theorem

$$\oint_C \mathbf{E} \cdot \mathrm{d}\mathbf{r} = -\frac{\partial}{\partial t}\left(\int_S \mathbf{H} \cdot \mathrm{d}\mathbf{S}\right)$$

which is a statement of Faraday's law of electromagnetic induction. Again, this is more commonly stated in the form that a current is induced in a wire C by a magnetic field flux through the surface enclosed by the loop of wire. The minus sign will be familiar to those that remember physics and the 'back emf' induced by a current flowing through a coil. The third of Maxwell's equations can be written as a volume integral as follows:

$$\int_V \nabla \cdot \mathbf{E}\mathrm{d}V = \int_V \rho \mathrm{d}V = Q$$

or, using the Divergence Theorem

$$\int_S \mathbf{E} \cdot \mathrm{d}\mathbf{S} = Q$$

This states that the flux of the electric field over the closed surface S is the same as the total charge inside S. This is called Gauss's Law and is a three-dimensional version of Ampère's circuit law for currents in a wire. Finally, the statement that $\nabla \cdot \mathbf{H} = 0$ is equivalent, via the Divergence Theorem to saying that:

$$\int_S \mathbf{H} \cdot \mathrm{d}\mathbf{S} = 0$$

that is that there are no isolated magnetic poles. Magnetic poles occur in pairs like small bar magnets, one south pole and one north pole.

In reality, we have worked backwards in this example. Physicists used experiment to deduce laws in terms of relationships between magnetic and electric fields in the early part of the nineteenth century, then along came James Clerk Maxwell late in the 1880s to formulate these laws into his famous four equations.

11.3 Exercises

11.1. Calculate the flux of the vector field $\mathbf{F} = (x - y^2)\mathbf{i} + 3y\mathbf{j} + (x^3 - z)\mathbf{k}$ in the following two cases:
(a) from the sphere $x^2 + y^2 + z^2 = 9$,
(b) from the cube with vertices $(0, 0, 0)$, $(a, 0, 0)$, $(0, a, 0)$, $(0, 0, a)$, $(a, a, 0)$, $(a, 0, a)$, $(0, a, a)$ and (a, a, a).
In each case, use the Divergence Theorem.

11.2. Verify the Divergence Theorem for $\mathbf{F} = \mathbf{r}$, and S is the surface of the cylinder $x^2 + y^2 \leq 1$, $0 \leq z \leq 2$.

11.3. Verify the Divergence Theorem for the vector field and surface of Example 10.7; $\mathbf{F} = x^2\mathbf{i} + y^2\mathbf{j} + z^2\mathbf{k}$, S is the sphere $x^2 + y^2 + z^2 = 1$.

11.4. Find the flux of the vector field $\mathbf{F} = y^2z^3\mathbf{i} + 4x^2yz^2\mathbf{j} + x^2z^3\mathbf{k}$ across the unit sphere $x^2 + y^2 + z^2 = 1$.

11.5. Use the Divergence Theorem to show that if $\mathbf{F} = \nabla\phi$ for some ϕ then ϕ is harmonic if and only if $\nabla \cdot \mathbf{F} = 0$.

11.6. Develop an alternative proof of Stokes' Theorem as follows. Write the surface S in the parametric form $x = u$, $y = v$, $z = f(u, v)$ and use Exercise 10.12 to show that $\int_S \mathbf{F} \cdot \mathrm{d}\mathbf{S} = \int_R \left(-F_1 \frac{\partial z}{\partial x} - F_2 \frac{\partial z}{\partial y} + F_3\right) \mathrm{d}x\mathrm{d}y$ where $\mathbf{F} = F_1\mathbf{i} + F_2\mathbf{j} + F_3\mathbf{k}$.
Now write $\mathbf{F} = \nabla \times \mathbf{A}$ to obtain $\int_S \nabla \times \mathbf{A} \cdot \mathrm{d}\mathbf{S}$ in Cartesian form.
Consider $\int_C \mathbf{A} \cdot \mathrm{d}\mathbf{s}$ and write this in the parametric form $\int_a^b \left(A_1 \frac{\mathrm{d}x}{\mathrm{d}t} + A_2 \frac{\mathrm{d}y}{\mathrm{d}t} + A_3 \frac{\mathrm{d}z}{\mathrm{d}t}\right) \mathrm{d}t$ where C is the boundary to the surface S and the parameterisation is $\mathbf{r} = \mathbf{r}(t)$. Now use Green's Theorem in the Plane (Example 9.8) to establish Stokes' Theorem. This is an entirely legitimate alternative proof to that of Example 11.2.

11.7. Use Gauss's Divergence Theorem to prove, for any scalar function ϕ with continuous derivatives inside and on the surrounding surface S of a volume V that

$$\int_V \nabla\phi \mathrm{d}V = \int_S \phi \mathrm{d}\mathbf{S}$$

This relationship is particularly useful in determining the equation of motion of a moving fluid.

11.8. A function ϕ is harmonic inside a volume V. Use the Divergence Theorem to show that any maxima or minima must be attained on the boundary of V.

11.9. Use the Divergence Theorem to show that, for a vector field $\mathbf{F}$ with continuous derivatives inside and on the closed surface S

$$\int_V \nabla \times \mathbf{F} \mathrm{d}V = -\int_S \mathbf{F} \times \mathrm{d}\mathbf{S}$$

where V is the volume enclosed by S.

11.10. Verify Stokes' Theorem for the following fields and surfaces:
(a) $\mathbf{F} = (2x - y)\mathbf{i} - yz^2\mathbf{j} - y^2z\mathbf{k}$, where S is the hemispherical surface $x^2 + y^2 + z^2 = a^2$, where a is a constant.
(b) $\mathbf{F} = -3y\mathbf{i} + 3x\mathbf{j} + z^4\mathbf{k}$, where S is that portion of the ellipsoid $2x^2 + 2y^2 + z^2 = 1$ that lies above the plane $z = 1/\sqrt{2}$.

11.11. The cylinder $x^2 + y^2 = b^2$ intersects the plane $y + z = a^2$ in a curve C. Assuming that $a^2 \geq b > 0$ calculate
$\oint_C (xy\mathbf{i} + yz\mathbf{j} + xz\mathbf{k}) \cdot \mathrm{d}\mathbf{r}$ by using Stokes' Theorem.

11.12. In Examples 10.8 and 11.12 Maxwell's equations were introduced. Slightly modified to include a current $\mathbf{J}$, but to exclude time variation they become:

$$\nabla \times \mathbf{H} = \mathbf{J}$$
$$\nabla \times \mathbf{E} = \mathbf{0}$$
$$\nabla \cdot \mathbf{E} = \rho$$
$$\nabla \cdot \mathbf{H} = 0$$

Re-introducing the Poynting vector $\mathbf{P} = \mathbf{E} \times \mathbf{H}$, if S is a closed surface, show that the radiation of $\mathbf{P}$ through the surface S is given by

$$\int_S \mathbf{P} \cdot \mathrm{d}\mathbf{S} = -\int_V \mathbf{E} \cdot \mathbf{J} \mathrm{d}V$$

where V is the volume enclosed by S.

Hints and Answers to Exercises

Exercises 1.3

1.1. (a) 11, (b) $-\frac{2}{7}$, (c) -4, (d) $\sqrt{5}$, (e) $\frac{1}{2}(b-a)$, (f) $\frac{3}{5}$, (g) $\frac{1}{7}$, (h) 1, (i) 1.

1.2. Terminal speed 8 m s^{-1}, 1.24 seconds.

1.3. Use $\dfrac{f(x+h)-f(x)}{h}$ and let $h \to 0$, so

(a) $\dfrac{(x+h)^2-x^2}{h} = \dfrac{x^2+2hx+h^2-x^2}{h} = 2x+h \to 2x$ as $h \to 0$.

(b) $\dfrac{\cos(x+h)-\cos x}{h}$

$= \dfrac{\cos x\cos h - \sin x \sin h - \cos x}{h}$

$\approx \dfrac{\cos x - h\sin x - \cos x}{h}$

for small h. So $\dfrac{\cos(x+h)-\cos x}{h} \to -\sin x$ as $h \to 0$.

(c) $\dfrac{e^{x+h}-e^x}{h} = \dfrac{e^xe^h-e^x}{h} = e^x\left[\dfrac{e^h-1}{h}\right] \to e^x$ as $h \to 0$ since $e^h \approx 1+h$ for small h.

(d) $\dfrac{\dfrac{1}{x+h}-\dfrac{1}{x}}{h} = \dfrac{x-x-h}{hx(x+h)} = -\dfrac{1}{x(x+h)} \to -\dfrac{1}{x^2}$ as $h \to 0$.

1.4. $\dfrac{d}{dx}(y^n) = \lim_{\Delta x\to 0}\left\{\dfrac{[y(x+\Delta x)]^n-[y(x)]^n}{\Delta x}\right\}$

$= \lim_{\Delta x\to 0}\left\{\dfrac{[y(x+\Delta x)]^n-[y(x)]^n}{\Delta y}\cdot\dfrac{\Delta y}{\Delta x}\right\}$

from which the result follows, provided Δy and Δx tend to zero jointly in a uniform manner, and all the limits exist.

1.5. For this function, the derivative $f'(x)$ has the definition

$$f'(x) = \begin{cases}\sin x & x<0\\ 2x & x\ge 0\end{cases}$$

This implies that $f'(x)$ is continuous at the origin as both $\sin x$ and $2x$ are zero there. The second derivative, however, is as follows:

$$f''(x) = \begin{cases}\cos x & x<0\\ 2 & x\ge 0\end{cases}$$

which is not continuous at the origin. Therefore $f''(x)$ does not exist at the origin.

1.6. (a) $\frac{5}{2}$, min.; (b) 0, max.; $\pm\frac{1}{2}\sqrt{6}$ both min.; (c) 0, min.; (d) $-\sqrt{2}$, min.; $\sqrt{2}$, max.

1.7. The function $\tan x$ is discontinuous at $x = \frac{\pi}{2}$, therefore the conditions of Rolle's Theorem are violated.

1.8. This is Rolle's Theorem in slightly more precise terms than given in Example 1.8. The explanation is similar.

1.9.

Rolle's Theorem:	If a particle returns to its initial position, there is a time when its speed is zero.
Mean Value Theorem:	The average speed is attained at least once by the object.

1.10. (a) $x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5$; (b) $1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4$; (c) $x - \frac{1}{2}x^2 + \frac{1}{3}x^3$; (d) $1 + \frac{1}{2}x - \frac{1}{8}x^2$.

1.11. The nth derivative of a polynomial is zero. The definition of a polynomial is $P(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ for some positive integer n. This is precisely the form of a truncated Maclaurin Series. A polynomial is thus its own Maclaurin Series, with

$a_k = \frac{1}{k!}P^{(k)}(0)$ for $k = 1, 2, \ldots, n$.

1.12. (a) $x\ln x = (x-2+2)\ln(x-2+2) = (u+2)\ln(u+2)$ where $u = x-2$. Expanding in terms of u then writing $u = x-2$ gives

$$x\ln x = 2\ln 2 + (1+\ln 2)(x-2) + \sum_{k=0}^{\infty}\frac{(-1)^k}{k(k-1)2^{k-1}}(x-2)^k$$

(b) $(1-2x)^{-3} = 5^{-3}\sum_{k=0}^{\infty}(k+2)(k+1)\dfrac{2^{k-1}}{5^k}(x+2)^k$

(write $u = x+2$).

(c) $\sin x = \sin(x-\pi+\pi) = -\sin(x-\pi)$ hence

$$\sin x = \sum_{k=0}^{\infty}\frac{(-1)^{k+1}(x-\pi)^{2k+1}}{(2k+1)!}$$

Note, always use the *easiest* method of expanding a function (by calculus, binomial theorem, using known series etc.). Theory then tells us that the series found is *the* Taylor's Series since it is unique.

1.13. $\sqrt{61}$ is evaluated most straightforwardly by considering $\sqrt{61} = \sqrt{64 - 3} = 8\left(1 - \frac{3}{64}\right)^{1/2}$ and expanding using the binomial theorem. Whence

$$\sqrt{61} = 8\left(1 - \frac{3}{128} - \frac{1}{8}\left(\frac{3}{64}\right)^2 - \frac{1}{16}\left(\frac{3}{64}\right)^3 - \ldots\right)$$

that is $\sqrt{61} = 7.81025$

1.14. (a) $\frac{8}{9}$, (b) 0, (c) $\frac{1}{9}$, (d) -1, (e) 0, (f) $-\frac{25}{49}$, (g) $a - b$, (h) $\ln a$ (recall that $\frac{d}{dx}(a^x) = a^x \ln a$ in using L'Hôpital's Rule), (i) $-\frac{1}{3}$ (it is necessary to use L'Hôpital's Rule three times here), (j) $\frac{1}{2}$ (in this example, use that if $\lim_{x \to L} f(x) = C$ then $\lim_{x \to L} \sqrt{f(x)} = \sqrt{C}$), (k) 0, (l) 1, (m) 1 (take logs and use that if $\lim_{x \to L} f(x) = C$ then $\lim_{x \to L}(\ln f(x)) = \ln(C)$), (n) 5 (L'Hôpital's Rule is not essential here, but take logs if you want to use it).

1.15. The Maclaurin Series for the two parts of the numerator are

$$\sin(\tan x) = x + \frac{1}{6}x^3 - \frac{1}{40}x^5 - \frac{55}{1008}x^7 + \ldots, \text{ and}$$

$$\tan(\sin x) = x + \frac{1}{6}x^3 - \frac{1}{40}x^5 - \frac{107}{5040}x^7 + \ldots$$

The first three terms of each are identical, so when they are subtracted we obtain $\sin(\tan x) - \tan(\sin x) = -\frac{x^7}{30} + \ldots$.

Hence the limit in the question has the value $-\frac{1}{30}$. It is a brave student who attempts this by hand; I used computer algebra.

1.16. Start with the calculator's best estimate as first guess. If you started with $\sqrt{2} = 1.4142136$, then the Newton–Raphson method applied to the equation $f(x) = x^2 - 2$ yields.

$$x_1 = 1.4142136 - \frac{(1.4142136)^2 - 2}{2 \times 1.4142136} = 1.41421356$$

an improvement. Similarly, for $\sqrt{7}$ and $\sqrt[3]{6}$ the start values are 2.6457513 and 1.8171206 and these are improved after one step of the Newton–Raphson method to 2.645751311 and 1.817120593 respectively.

1.17. With $f(x) = x^2 - a$, $f'(x) = 2x$ so the Newton–Raphson method yields

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{x_n^2 - a}{2x_n} \text{ which gives the result.}$$

1.18. The integral is equal to the limit of the function $\frac{1}{n}\sum_{k=1}^{\infty}\frac{k^3}{n^3}$ as $n \to \infty$ which, using the summation given leads to the limit

$$\lim_{n\to\infty}\left\{\frac{1}{n}\frac{n^2}{4}\frac{1}{n^3}(n + 1)^2\right\} = \frac{1}{4}.$$

1.19. (a) $\frac{\pi}{4} - \frac{1}{2}\ln 2$, (b) $\frac{1}{2}x^2\ln x - \frac{1}{4}x^2 + C$,

(c) $-\frac{xe^x}{x + 1} + e^x + C$ (choose $v = \frac{1}{(x + 1)^2}$, $u = xe^x$).

1.20. $I_n = \int \cos^n x dx$, then $I_n = \frac{1}{n}\cos^{n-1}x \sin x + \frac{n - 1}{n} I_{n-2}$, also

$I_4 = \frac{1}{4}\cos^3 x \sin x + \frac{3}{8}\cos x \sin x + \frac{3}{8}x + C$, and

$I_5 = \frac{1}{5}\cos^4 x \sin x + \frac{4}{15}\cos^2 x \sin x + \frac{8}{15}\sin x + C$.

1.21. $I_n = \int \sin^n x dx$, then $I_n = -\frac{1}{n}\sin^{n-1}x \cos x + \frac{n - 1}{n} I_{n-2}$.

$I_3 = -\frac{1}{3}\sin^2 x \cos x - \frac{2}{3}\cos x$ and

$I_5 = -\frac{1}{5}\sin^4 x \cos x - \frac{4}{15}\sin^2 x \cos x - \frac{8}{15}\cos x + C$.

1.22. (a) $\frac{1}{2}$, (b) $\ln\left(\frac{3}{5}\right)$, (c) $\frac{1}{2}$, (d) -2, (e) $\frac{\pi}{2ab}$.

1.23. $\frac{16}{3}\pi$.

1.24. $\frac{3}{5}h$.

Exercises 2.3

2.1. $f_x = 4x^3 + 15x^2y + 5y^3$, $f_y = 5x^3 + 15xy^2 + 4y^3$,
$f_{xx} = 12x^2 + 30xy$, $f_{yy} = 30xy + 12y^2$, $f_{xy} = 15(x^2 + y^2)$.

2.2. $f_x = (1 + x + y + z)e^{(x+y-z)}$, $f_y = (1 + x + y + z)e^{(x+y-z)}$,
$f_z = (1 - x - y - z)e^{(x+y-z)}$, $f_{xy} = (2 + x + y + z)e^{(x+y-z)}$,
$f_{zx} = -(x + y + z)e^{(x+y-z)}$, $f_{yz} = -(x + y + z)e^{(x+y-z)}$,
$f_{xx} = (2 + x + y + z)e^{(x+y-z)}$, $f_{yy} = (2 + x + y + z)e^{(x+y-z)}$,
$f_{zz} = (x + y + z - 2)e^{(x+y-z)}$.

2.3. (a) $\frac{\partial u}{\partial x} = 2x = \frac{\partial v}{\partial y}$, $\frac{\partial u}{\partial y} = -2y = -\frac{\partial v}{\partial x}$;

(b) $\frac{\partial u}{\partial x} = e^x \cos y = \frac{\partial v}{\partial y}$, $\frac{\partial u}{\partial y} = -e^x \sin y = -\frac{\partial v}{\partial x}$;

(c) $\frac{\partial u}{\partial x} = \frac{x}{x^2 + y^2} = \frac{\partial v}{\partial y}$, $\frac{\partial u}{\partial y} = \frac{y}{x^2 + y^2} = -\frac{\partial v}{\partial x}$;

(d) $\frac{\partial u}{\partial x} = \frac{y^2 - x^2}{(x^2 + y^2)^2} = \frac{\partial v}{\partial y}$, $\frac{\partial u}{\partial y} = \frac{-2xy}{(x^2 + y^2)^2} = -\frac{\partial v}{\partial x}$.

That each u and v above is also harmonic is easily checked.

2.4. $\frac{\partial f}{\partial x} = 2Bxy^2 + 4Ax^3$, $\frac{\partial f}{\partial y} = 2Bx^2y + 4Cy^3$, hence

$x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} = 2Bx^2y^2 + 4Ax^4 + 2Bx^2y^2 + 4Cy^4 = 4f$

as required.

2.5. $\frac{\partial x}{\partial R} = \cos\theta$, $\frac{\partial y}{\partial \theta} = R\sin\theta$, $\frac{\partial x}{\partial \theta} = -R\sin\theta$, $\frac{\partial y}{\partial R} = \sin\theta$, thus

$\frac{\partial x}{\partial R}\frac{\partial y}{\partial \theta} - \frac{\partial x}{\partial \theta}\frac{\partial y}{\partial R} = R\cos^2\theta + R\sin^2\theta = R$.

2.6. $\frac{\partial P}{\partial V} = -k\frac{T}{V^2}$, so $V\frac{\partial P}{\partial V} = -k\frac{T}{V} = -P$. Also $\frac{\partial P}{\partial T} = \frac{k}{V}$ therefore $T\frac{\partial P}{\partial T} = k\frac{T}{V}$ hence $V\frac{\partial P}{\partial V} + T\frac{\partial P}{\partial T} = 0$.

2.7. Using $\frac{dz}{dt} = \frac{\partial z}{\partial y}\frac{dy}{dt} + \frac{\partial z}{\partial x}\frac{dx}{dt}$ we get

$\frac{dz}{dt} = \frac{2x}{(x-y)^2}\cos t + \frac{2y}{(x-y)^2}\sin t$

which is $\frac{dz}{dt} = \frac{2}{(x-y)^2}$. Using direct differentation, we get

$\frac{dz}{dt} = \frac{d}{dt}\left(\frac{\cos t + \sin t}{\cos t - \sin t}\right) = \frac{2}{(\cos t - \sin t)^2}$ after using the quotient rule and a little trigonometry. The results are the same.

2.8. Differentiating the given function, $f_x = \frac{y^2}{(x+y)^2}$ and $f_y = \frac{x^2}{(x+y)^2}$. Differentiating again gives $f_{xx} = -\frac{2y^2}{(x+y)^3}$, $f_{yy} = -\frac{2x^2}{(x+y)^3}$ and $f_{xy} = \frac{2xy}{(x+y)^3}$.Thus, $x^2 f_{xx} + 2xyf_{xy} + y^2 f_{yy} = \frac{-2x^2y^2 + 2xy\cdot 2xy - 2x^2y^2}{(x+y)^3}$

$= 0.$

2.9. With $\phi = F\left(\frac{y}{x}\right)$, $\frac{\partial\phi}{\partial x} = -\frac{y}{x^2}F$ and $\frac{\partial\phi}{\partial x} = \frac{y}{x}F'$ from which the result follows.

2.10. Writing $x = \xi + \eta$ and $t = -\frac{1}{c}\eta$, using the chain rule gives

$\frac{\partial}{\partial\xi} = \frac{\partial}{\partial x}, \frac{\partial}{\partial\eta} = c\frac{\partial}{\partial x} - \frac{\partial}{\partial t}$, whence the given equation

$c\frac{\partial^2\phi}{\partial x^2} - \frac{\partial^2\phi}{\partial x\partial t} = 0$ implies $\frac{\partial^2\phi}{\partial\xi\partial\eta} = 0$ so that

$\phi = f(\xi) + g_0(\eta) = f(x + ct) + g(t).$

2.11. Differentiating $r^2 = x_1^2 + x_2^2 + \ldots + x_n^2$ yields $2r\frac{\partial r}{\partial x_i} = 2x_i$ for all $i = 1, 2, \cdots, n$. Thus $\sum_{i=1}^{n} x_i\frac{\partial r}{\partial x_i} = \sum_{i=1}^{n}\frac{x_i^2}{r} = \frac{r^2}{r} = r.$

2.12. $\frac{\partial w}{\partial x} = -\frac{wyz\cos xyz}{2w + \sin xyz}, \frac{\partial w}{\partial y} = -\frac{wxz\cos xyz}{2w + \sin xyz}$,

$\frac{\partial w}{\partial z} = -\frac{wxy\cos xyz}{2w + \sin xyz}$.

2.13. Using the chain rule, $\frac{\partial f}{\partial u} = e^u\sec v\frac{\partial f}{\partial x} + e^u\tan v\frac{\partial f}{\partial y}$ and

$\frac{\partial f}{\partial u} = e^u\sec v\tan v\frac{\partial f}{\partial x} + e^u\sec^2 v\frac{\partial f}{\partial y}$

Hence $\left(\frac{\partial f}{\partial u}\right)^2 + \cos^2 v\left(\frac{\partial f}{\partial v}\right)^2 =$

$(\sec^2 v - \tan^2 v)\left\{\left(\frac{\partial f}{\partial x}\right)^2 - \left(\frac{\partial f}{\partial y}\right)^2\right\}e^{2u}$ from which the result follows.

2.14. This exercise is quite difficult. Use the chain rule and keep a clear head!

$\frac{\partial\phi}{\partial x} = \frac{\partial\phi}{\partial u} + y\frac{\partial\phi}{\partial v}$ and $\frac{\partial\phi}{\partial y} = \frac{\partial\phi}{\partial u} + x\frac{\partial\phi}{\partial v}$ from which

$x\frac{\partial\phi}{\partial x} - y\frac{\partial\phi}{\partial y} = (x - y)\frac{\partial\phi}{\partial u}$ (eliminating $\frac{\partial\phi}{\partial v}$).

Expanding the operator $\left(x\frac{\partial}{\partial x} - y\frac{\partial}{\partial y}\right)\left(x\frac{\partial\phi}{\partial x} - y\frac{\partial\phi}{\partial y}\right) = \left(x\frac{\partial}{\partial x} - y\frac{\partial}{\partial y}\right)\left\{(x-y)\frac{\partial\phi}{\partial u}\right\}$ gives, taking care with the products:

$x^2\phi_{xx} - 2xy\phi_{xy} + y^2\phi_{yy} + x\phi_x + y\phi_y = x\phi_u + y\phi_u + (x-y)^2\phi_{uu}$ using the suffix derivative notation. However, using the first two results:

$x\frac{\partial\phi}{\partial x} + y\frac{\partial\phi}{\partial y} = (x+y)\frac{\partial\phi}{\partial u} + 2xy\frac{\partial\phi}{\partial v}$,

and $\frac{\partial\phi}{\partial x} - \frac{\partial\phi}{\partial y} = (x-y)\frac{\partial\phi}{\partial v}$. Thus

$$\begin{aligned} x^2\phi_{xx} - 2xy\phi_{xy} + y^2\phi_{yy} &= -(x+y)\phi_u - 2xy\phi_v + (x+y)\phi_u + (x-y)^2\phi_{uu} \\ &= -2xy\phi_v + (x-y)^2\phi_{uu} \end{aligned}$$

Multiplying this by $(x - y)$ yields

$$\begin{aligned} (x-y)(x^2\phi_{xx} - 2xy\phi_{xy} + y^2\phi_{yy}) &= -2xy(x-y)\phi_v + (x-y)^3\phi_{uu} \\ &= 2xy(\phi_x - \phi_y) + (x-y)^3\phi_{uu} \end{aligned}$$

hence if ϕ obeys the equation given in the question, we must have, $(x - y)^3\phi_{uu} = 0$. Integrating this twice with respect to u yields $\phi = (x + y)f(xy) + g(xy)$ where f and g are arbitrary functions.

2.15. The Jacobian $\frac{\partial(u,v)}{\partial(x,y)}$ is given by $\begin{vmatrix} \frac{1+y^2}{(1-xy)^2} & \frac{1+x^2}{(1-xy)^2} \\ \frac{1}{1+x^2} & \frac{1}{1+y^2} \end{vmatrix} = 0$

hence there is a functional relationship between u and v. To see what this is, write $\theta = \tan^{-1}x$, $\phi = \tan^{-1}y$, then $u = \theta + \phi$, and

$\tan u = \frac{\tan\theta + \tan\phi}{1 - \tan\theta\tan\phi} = \frac{x+y}{1-xy} = v$, so $v = \tan u$.

2.16. From $V = \frac{1}{3}\pi h(b^2 + ba + a^2)$ we get by logarithmic differentiation:

$$\frac{dV}{V} = \frac{dh}{h} + \frac{2 + \frac{b}{a}}{1 + \frac{b}{a} + \left(\frac{b}{a}\right)^2}\frac{da}{a} + \frac{2 + \frac{a}{b}}{1 + \frac{a}{b} + \left(\frac{a}{b}\right)^2}\frac{db}{b}$$

and using the numbers in the question gives

$\frac{dV}{V} = \frac{1}{100}\left(3 + \frac{8}{9} + \frac{10}{7}\right) = 5.32$ per cent.

2.17. $A = \frac{1}{2}ab\sin C$, whence $\frac{dA}{A} = \frac{da}{a} + \frac{db}{b} + \cot C dC$ from which the maximum error is 12.7 per cent. In this example, if C is very small the area is very sensitive to changes in its value. This is very important in numerical analysis where it is related to ill-conditioning.

2.18. $\int_0^\infty xe^{-ax}\cos bx\,dx = \frac{a^2 - b^2}{(a^2+b^2)^2}, \int_0^\infty xe^{-ax}\sin bx\,dx = \frac{2ab}{(a^2+b^2)^2}.$

2.19. $\int_0^\infty t^n e^{-st}dt = \frac{n!}{s^{n+1}}$ for all positive integers n.

Exercises 3.3

3.1. (a) $e^{xy} = 1 + xy + \frac{x^2y^2}{2!} + \ldots$ by direct expansion. The question only asks for second-order terms, therefore $e^{xy} \approx 1 + xy$ suffices. About the point (0, 1) we have

$$e^{xy} = 1 + x + x(y - 1) + \tfrac{1}{2}x^2 + O(x^3, (y - 1)^3),$$

and about the point (1, 0) we have

$$e^{xy} = 1 + y + y(x - 1) + \tfrac{1}{2}y^2 + O((x - 1)^3, y^3).$$

Finally, about the point (1, 1) we have

$$e^{xy} = e + e(x - 1) + e(y - 1) + \tfrac{1}{2}e[(y - 1)^2 + 4(x - 1)(y - 1) + (x - 1)^2] + \cdots$$

(b) Using Taylor's Series directly gives, about (1, 0),

$$\operatorname{cosec}(x + y) = \operatorname{cosec}(1) - ((x - 1) + y)\operatorname{cosec}(1)\cot(1) + \tfrac{1}{2}((x - 1)^2 + 2(x - 1)y + y^2)(\operatorname{cosec}(1)\cot^2(1) + \operatorname{cosec}^3(1)) + \cdots$$

and about (0, 1),

$$\operatorname{cosec}(x + y) = \operatorname{cosec}(1) - (x + (y - 1))\operatorname{cosec}(1)\cot(1) + \tfrac{1}{2}(x^2 + 2x(y - 1) + (y - 1)^2)(\operatorname{cosec}(1)\cot^2(1) + \operatorname{cosec}^3(1)) + \cdots$$

3.2. $$\frac{1}{2 + x_1x_2x_3x_4} = \frac{1}{2} - \frac{1}{4}x_1x_2x_3x_4 + \frac{1}{8}(x_1x_2x_3x_4)^2 + \cdots$$

3.3. $$\frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} + \frac{1}{x_4} = \frac{1}{u_1} + \frac{1}{u_2} + \frac{1}{u_3} + \frac{1}{u_4} - \frac{h_1}{u_1^2} - \frac{h_2}{u_2^2} - \frac{h_3}{u_3^2} - \frac{h_4}{u_4^2} + \frac{h_1^2}{u_1^3} + \frac{h_2^2}{u_2^3} + \frac{h_3^2}{u_3^3} + \frac{h_4^2}{u_4^3} + \cdots$$

where the h's are the variables.

3.4. (a) (2, 6) minimum,
(b) (0, 0) saddle, (1, 1), (−1, −1) both maxima.
(c) $f = -xye^{-\frac{1}{2}(x^2+y^2)}$, so $f_x = y(x^2 - 1)e^{-\frac{1}{2}(x^2+y^2)}$,
$f_y = x(y^2 - 1)e^{-\frac{1}{2}(x^2+y^2)}$, $f_{xx} = xy(3 - x^2)e^{-\frac{1}{2}(x^2+y^2)}$,
$f_{yy} = xy(3 - y^2)e^{-\frac{1}{2}(x^2+y^2)}$, and $f_{xy} = (x^2 - 1)(1 - y^2)e^{-\frac{1}{2}(x^2+y^2)}$.
This leads to the following extrema:
(0, 0) saddle, (1, 1) minimum, (−1, 1) maximum,
(1, −1) maximum, (−1, −1) minimum.
(d) Extremum whenever $x = 2y$ ($x \neq 0$).
By writing $x = 2t + \varepsilon\cos\theta$, $y = \frac{1}{t} + \varepsilon\sin\theta$ and expanding the function $f(x, y) = -\frac{x^2 + 2y^2}{(x + y)^2}$ in powers of ε, we obtain:

$$f\left(2t + \varepsilon\cos\theta, \frac{1}{t} + \varepsilon\sin\theta\right) - f\left(2t, \frac{1}{t}\right) = -\frac{\varepsilon^2}{9t^2}\left[\cos^2\theta + 2\sin^2\theta + \frac{10}{3}(\cos\theta + \sin\theta)^2\right]$$

and this is < 0 for all θ. Hence $x = 2y$ is a line of maxima.
(e) $f(x, y) = e^x\cos y$ has no extreme values.

3.5. $f(x, y) = \sin xy$, so $f_x = y\cos xy$, $f_y = x\cos xy$. Lines of extrema along rectangular hyperbolae $xy = c$ where $c = \frac{2n + 1}{2}\pi$, n an integer. Using the parameterisation $x = t + \varepsilon\cos\theta$, $y = \frac{1}{t} + \varepsilon\sin\theta$ it is shown that

$$f(x, y) = \sin xy \text{ has maxima at } xy = \frac{\pi}{2}, \frac{5\pi}{2}, \frac{9\pi}{2}, \ldots \text{ and}$$

$$f(x, y) = \sin xy \text{ has minima at } xy = \frac{3\pi}{2}, \frac{7\pi}{2}, \frac{11\pi}{2} \ldots.$$

3.6. If $f(x, y) = 4 - \sqrt{x^2 + y^2}$ then, by inspection, 4 is a maximum value occurring at the origin (0, 0). This function represents a cone with the vertex at the origin at which derivatives do not exist, therefore normal tests for extrema fail.

3.7. $f(x, y) = \sin x \sin y$ has extrema wherever either $x = \frac{2k + 1}{2}\pi$, $y = \frac{2l + 1}{2}\pi$, or $x = r\pi$, $y = s\pi$ where k, l, r, s are integers. This leads to the following classification: $x = \frac{2k + 1}{2}\pi$, $y = \frac{2l + 1}{2}\pi$ is a maximum if $(k + l)$ is odd and a minimum if $(k + l)$ is even, $x = r\pi$, $y = s\pi$ is always a saddle point for all r and s.

3.8. $f(x, y) = x^4 - 2(x - y)^2 + y^4$ leads to extrema wherever the following two equations hold: $x^3 = x - y$ and $y^3 = y - x$, hence the three turning points are (0, 0), $(\sqrt{2}, -\sqrt{2})$ and $(-\sqrt{2}, \sqrt{2})$. Using the usual procedure, it is straightforward to deduce that $(\sqrt{2}, -\sqrt{2})$ and $(-\sqrt{2}, \sqrt{2})$ are both minima. However, at the origin the second derivatives of $f(x, y)$, f_{xx}, f_{yy} and f_{xy} all have the value 4 so $f_{xx}f_{yy} = f_{xy}^2$. However, near the origin, it is quite easy to deduce that $f \approx -2\varepsilon^2(\cos\theta - \sin\theta)^2 < 0$ by writing $x = \varepsilon\cos\theta$, $y = \varepsilon\sin\theta$. Hence the origin is a maximum.

3.9. Extreme values are at the five points (0, 0), (α, α), $(-\alpha, \alpha)$, $(\alpha, -\alpha)$ and $(-\alpha, -\alpha)$. The last four are all saddle points, since $f_{xx} = f_{yy} = 0$ yet $f_{xy}^2 = \sin^2 x \sin^2 y > 0$. At the origin, all the second derivatives of f are zero, so we put $x = \varepsilon\cos\theta$, $y = \varepsilon\sin\theta$ to obtain the expansion

$f(\varepsilon\cos\theta, \varepsilon\sin\theta) = 1 + \frac{\varepsilon^4}{24}\cos 4\theta + O(\varepsilon^6)$, whence (0, 0) is a saddle point too.

3.10. With $f(x, y, z) = xyz(1 - x - y - z)$, the first derivatives are zero under the following conditions:

$$1 - 2x - y - z = 0 \text{ or } yz = 0$$
$$1 - x - 2y - z = 0 \text{ or } xz = 0$$
$$1 - x - y - 2z = 0 \text{ or } yx = 0$$

If either x, or y or z is zero then either $y + z = 1$, or $x + z = 1$ or $y + x = 1$. In any of these cases, $f(x, y, z) = xyz(1 - x - y - z) = 0$. On the other hand, if none of x, y, z is zero, the only solution of the three conditions is $x = y = z = \frac{1}{4}$ in which case

$$f(x, y, z) = xyz(1 - x - y - z) = \frac{1}{4}\cdot\frac{1}{4}\cdot\frac{1}{4}\left(1 - \frac{1}{4} - \frac{1}{4} - \frac{1}{4}\right) = \frac{1}{256},$$

which is obviously a maximum.

3.11. $x = 1, y = 0$ gives $f = 2$, a minimum. $x = 0, y = 1$ gives $f = 4$, a maximum.

3.12. The equations $\frac{4}{3}\pi bc + \lambda = 0, \frac{4}{3}\pi ac + \lambda = 0,$ $\frac{4}{3}\pi ab + \lambda = 0$ result from differentiating the Lagrangian $H = \frac{4}{3}\pi abc + \lambda(a + b + c - 1)$ and putting each derivative equal to zero. These imply $a = b = c$, that is the ellipsoid is a sphere.

3.13. With $H = \pi r^2 h + \lambda(2\pi r^2 + 2\pi rh - 1)$, differentiation gives $r = 2h$, whence $h = \frac{1}{\sqrt{12\pi}}, r = \frac{1}{\sqrt{3\pi}}$. This gives a volume of $\frac{1}{6\sqrt{3\pi}}$, obviously a maximum.

3.14. With the box (cuboid) with dimensions $2x$, $2y$ and z the Lagrangian is $H = 4xyz + \lambda(z + \sqrt{x^2 + y^2} - 9)$ from which $y = x, z = \frac{x}{\sqrt{2}}$. The equation of the cone then implies that $z + \sqrt{x^2 + y^2} - 9 = 0 \Rightarrow z + x\sqrt{2} = 9$ so that $x = \dfrac{9}{\sqrt{2} + \frac{1}{\sqrt{2}}} = 3\sqrt{2} = y, z = 3$. Thus the dimensions are $3\sqrt{2}$, $3\sqrt{2}$ and 3, giving a volume of $V = 216$.

3.15. The Lagrangian for this problem is

$$H = \frac{\sqrt{d_1^2 + x^2}}{v_1} + \frac{\sqrt{d_2^2 + y^2}}{v_2} + \lambda(x + y - a).$$

It is essential to eliminate the angles as these depend on x and y. The conditions

$$\frac{\partial H}{\partial x} = \frac{1}{2}(d_1^2 + x^2)^{-1/2} \cdot \frac{2x}{v_1} + \lambda = 0$$

$$\frac{\partial H}{\partial x} = \frac{1}{2}(d_2^2 + y^2)^{-1/2} \cdot \frac{2x}{v_2} + \lambda = 0 \text{ hold,}$$

from which $\dfrac{x}{v_1\sqrt{d_1^2 + x^2}} = \dfrac{y}{v^2\sqrt{d_2^2 + y^2}}$ or $\dfrac{\sin\theta_1}{v_1} = \dfrac{\sin\theta_2}{v_2}$.

3.16. If the corridor has width a and b then the length of the rod is $l = b\operatorname{cosec}\theta + a\sec\theta$ so $\frac{\partial l}{\partial \theta} = 0$ implies $\tan^3\theta = b/a$. Hence $l = (a^{2/3} + b^{2/3})^{3/2}$.

Exercises 4.3

4.1. $f(x, y) = 4x^2 - 4xy + 2y^2$, so $\nabla f = \begin{pmatrix} 8x - 4y \\ -4x + 4y \end{pmatrix}$. The line direction is $\mathbf{x}_1 = \begin{pmatrix} 2 - 4\lambda \\ 3 - 4\lambda \end{pmatrix}$. Evaluating ∇f on this line then demanding that $\mathbf{x}_1 \cdot \nabla f$ is minimal gives $-32 + 64\lambda = 0 \Rightarrow \lambda = \frac{1}{2}$ whence $\mathbf{x}_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Repeating this process gives $\mathbf{x}_2 = \begin{pmatrix} 4\lambda \\ 1 - 4\lambda \end{pmatrix}$ and $\lambda = \frac{1}{10}$, so the point $\begin{pmatrix} 0.4 \\ 0.6 \end{pmatrix}$ is the result of two iterations.

4.2.
$$f(x,y) = -(x - 1)^4 - (x - y)^2,$$
$$\text{so } \nabla f = \begin{pmatrix} -4(x - 1)^3 - 2(x - y) \\ 2(x - y) \end{pmatrix}$$

Starting the first iteration with (0, 0), we get to (0.41025, 0) with $\lambda = 0.10256$ after the second iteration. The third iteration leads to (0.41025, 0.41025) via $\lambda = 0.5$.

4.3. Putting $\mathbf{g} = \nabla f$ in the Taylor Polynomial gives the result.

4.4. $f(x, y) = x^4 - 2xy + (y + 2)^2$, starting at $\left(\frac{1}{2}, 1\right)$ leads to the following formulae:

$\nabla f = \begin{pmatrix} 4x^3 - 2y \\ 2y - 2x + 4 \end{pmatrix}$, $\mathbf{G} = \begin{pmatrix} 12x^2 & -2 \\ 0 & 2 \end{pmatrix}$. At $\left(\frac{1}{2}, 1\right)$ we therefore have $\nabla f = \begin{pmatrix} -\frac{3}{2} \\ 5 \end{pmatrix}$, $\mathbf{G} = \begin{pmatrix} 3 & -2 \\ 0 & 2 \end{pmatrix}$

and $\mathbf{G}^{-1} = \frac{1}{6}\begin{pmatrix} 2 & 2 \\ 0 & 3 \end{pmatrix}$, whence $\mathbf{x}_1 = \left(4, \frac{5}{2}\right)$.

This means that the Newton–Raphson method has actually taken us further *away* from the actual minimum. There is therefore little point in carrying on with the method (see the answer to the next exercise).

4.5. Using software, the minimum value of $f(x, y) = x^4 - 2xy + (y + 2)^2$ is $x = -1.165373$, $y = -3.165373$

4.6. $f(x, y) = x^4 + xy + y^2 + 2y$, the Newton–Raphson method gives the results:

$$\mathbf{x}_0 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}, \mathbf{x}_1 = \begin{pmatrix} 0.7826 \\ -1.3913 \end{pmatrix}, \mathbf{x}_2 = \begin{pmatrix} 0.7058 \\ -1.3529 \end{pmatrix}, \mathbf{x}_3 = \begin{pmatrix} 0.6960 \\ -1.3480 \end{pmatrix}$$

4.7. $f(x, y) = x^2 + 8y^2 + z^2 + 2xy + 4yz + 8y - 2z$

$\nabla f = \begin{pmatrix} 2x + 2y \\ 2x + 16y + 4z + 2 \\ 4y + 2z - 2 \end{pmatrix}$ and $\mathbf{G} = \begin{pmatrix} 2 & 2 & 0 \\ 2 & 16 & 4 \\ 0 & 4 & 2 \end{pmatrix}$ which is of course constant.

Inverting $\mathbf{G}$ gives $\mathbf{G}^{-1} = \frac{1}{6}\begin{pmatrix} 4 & -1 & 2 \\ -1 & 1 & -2 \\ 2 & -2 & 7 \end{pmatrix}$, whence

$\mathbf{x}_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} - \frac{1}{6}\begin{pmatrix} 4 & -1 & 2 \\ -1 & 1 & -2 \\ 2 & -2 & 7 \end{pmatrix}\begin{pmatrix} 0 \\ 2 \\ -2 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix}$ which is the *exact* minimum.

4.8. Using the DFP method gives $\begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$. Using the method of steepest descent gives

$$\nabla f = \begin{pmatrix} 2(x - y) + \frac{1}{4}(x + y + 1)^3 \\ -2(x - y) + \frac{1}{4}(x + y + 1)^3 \end{pmatrix}.$$

At $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $\nabla f = \begin{pmatrix} \frac{1}{4} \\ \frac{1}{4} \end{pmatrix}$, $\mathbf{x}_1 = \begin{pmatrix} -\frac{1}{4}\lambda \\ -\frac{1}{4}\lambda \end{pmatrix}$

so that on $\mathbf{x}_1$, $\nabla f = \frac{1}{4}\begin{pmatrix} (1 - \frac{1}{2}\lambda)^3 \\ (1 - \frac{1}{2}\lambda)^3 \end{pmatrix}$. Demanding that $\mathbf{x}_1 \cdot \nabla f = 0$ implies $\lambda = \frac{1}{2}$ whence $\mathbf{x}_1 = \begin{pmatrix} -\frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$, the exact result. Using the Newton–Raphson method gives

$\begin{pmatrix}-0.1667\\-0.1667\end{pmatrix}$ and $f = 0.0123$ after one iteration, and $\begin{pmatrix}-0.2778\\-0.2778\end{pmatrix}$ and $f = 0.0024$ after two.

4.9. A start vector that makes the residual $\mathbf{r}$ as close to $\mathbf{0}$ as possible is the choice $a = 4$ and $4e^b = 2$ so $b = \ln(0.5) \approx -1$. So choose $(a, b) = (4, -1)$. The DFP method then gives the following results:

Step	a	b	$\mathbf{r}^T\mathbf{r}$
0	4	−1	1.1773
1	4.0748	−0.5652	0.1779
2	3.9388	−0.5219	0.1490
3	3.8903	−0.5349	0.1401
4	3.8884	−0.5349	0.1401
5	3.8884	−0.5349	0.1401

that is, convergence after 4 steps.

4.10. Denoting D_2 by (0, 0) and S by (x, y) then D_1 is (0, 6) and D_3 is (4, 0). We thus minimise the function $f(x, y) = \sqrt{x^2 + y^2} + \sqrt{(x - 4)^2 + y^2} + \sqrt{x^2 + (y - 6)^2}$. The numerical procedure is set up as follows:
First of all, we calculate ∇f, then set $\mathbf{H}_0 = \mathbf{I}$ and $\mathbf{y}_{i+1} = \nabla f_{i+1} - \nabla f_i$. For this problem

$$\nabla f = \begin{pmatrix} x(x^2 + y^2)^{-1/2} + (x - 4)((x - 4)^2 + y^2)^{-1/2} \\ \quad + x(x^2 + (y - 6)^2)^{-1/2} \\ y(x^2 + y^2)^{-1/2} + y((x - 4)^2 + y^2)^{-1/2} \\ \quad + (y - 6)(x^2 + (y - 6)^2)^{-1/2} \end{pmatrix}$$

The step length $\mathbf{h}$ is our decision, whence $\mathbf{x}_1 = \mathbf{h} - \mathbf{x}_0$. We have $\nabla f_0 = \begin{pmatrix}-1\\-1\end{pmatrix}$ thus we can find $\mathbf{y}_1 = \nabla f_1 - \nabla f_0$ (after finding ∇f_1 from $\mathbf{x}_1 = \mathbf{h} - \mathbf{x}_0$). Then we either use the DFP algorithm

$$\mathbf{H}_{i+1} = \mathbf{H}_i - \frac{\mathbf{H}_i\mathbf{y}_i\mathbf{y}_i^T\mathbf{H}_i}{\mathbf{y}_i^T\mathbf{H}_i\mathbf{y}_i} + \frac{\mathbf{h}_i\mathbf{h}_i^T}{\mathbf{h}_i^T\mathbf{y}_i}$$

or the more complex BFGS algorithm

$$\mathbf{H}_{i+1} = \mathbf{H}_i + \left(1 + \frac{\mathbf{y}_i^T\mathbf{y}_i}{\mathbf{h}_i^T\mathbf{y}_i}\right)\frac{\mathbf{h}_i\mathbf{h}_i^T}{\mathbf{h}_i^T\mathbf{y}_i} - \frac{\mathbf{h}_i\mathbf{y}_i^T + \mathbf{y}_i\mathbf{h}_i^T}{\mathbf{h}_i^T\mathbf{y}_i}$$

with $i = 1, 2, \ldots$ the right-hand side of each being known for $i = 1$. The solution to this soakaway location problem is (1.06566, 0.95725).

4.11. The DFP algorithm works reasonably well when applied to Rosenbrock's function $100(y - x^2)^2 + (1 - x)^2$. Steepest descent and Newton–Raphson fail to converge because of the curved shape of the valley in which the minimum lies. The 12th, 13th and 14th steps give the co-ordinates

0.9971091	0.9941688
0.9999877	0.9999816
1.0000016	1.0000031

obviously very close to (1, 1).

4.12. (a) The Hessian matrix for Powell's function is singular, therefore *no* method works well, even though (0, 0, 0, 0) is an obvious minimum.
(b) Fletcher and Powell's helix function has a helical-shaped valley leading to the point (1, 0, 0).
(c) Wood's function presents problems because there is a saddle very close to the minimum at (1, 1, 1, 1).

4.13. With $L(x, y, \lambda) = x^2 + y^2 + \lambda(xy - 3)$, zero partial derivatives of L soon reveal that $\lambda = 2$, $x = y = \sqrt{3}$ and $f = 6$.

4.14. The best way to tackle this problem is to calculate ∇f, evaluate it at the point and then check that this point is stationary by computing certain gradients. First of all, it is easy to check that the point $\left(\frac{1}{2}, \frac{1}{3}, \frac{7}{4}\right)$ lies on both constraints. Then if we define $\nabla\mathbf{c} = \begin{pmatrix} 2x & 1 \\ 0 & 3 \\ 1 & 2 \end{pmatrix}$ where each column consists of the three partial derivatives of the two constraints, then a *feasible direction* is $\mathbf{s}$ where $\mathbf{s} \cdot \nabla\mathbf{c} = (0, 0)$. To find $\mathbf{s}$ takes us into the practice of finding eigenvectors (see the Fact Sheet, section 4.1). At the point $\left(\frac{1}{2}, \frac{1}{3}, \frac{7}{4}\right)$, $\mathbf{s} = k(1, \frac{1}{3} -1)$ where k is any constant. It is now easy to check that $\nabla f = \begin{pmatrix} \frac{1}{2} \\ -3 \\ -\frac{1}{2} \end{pmatrix}$ satisfies $\mathbf{s} \cdot \nabla f = 0$. This means that the point $\left(\frac{1}{2}, \frac{1}{3}, \frac{7}{4}\right)$ is indeed an extremum. It is in fact a minimum, which is most easily verified from first principles.

4.15. The equivalent equality constraint model is:

$$\begin{aligned} \text{minimise } f(\mathbf{x}) &= x^2 - 2xy + 2y^2 + 6x + 7y \\ \text{subject to: } 25 - x^2 - y^2 - u^2 &= 0 \\ -x - 3y - v^2 &= 0 \end{aligned}$$

where u^2 and v^2 are surplus variables. Using software, the solution to this problem is $(x, y) = (-3, -4)$. [The topic of inequality constraints is a great deal more complex than implied here, but is well outside the scope of this Work Out.]

4.16. Solutions are $(x, y, z) = (0.4, 1, -0.6)$. The Newton–Raphson method is exact for quadratic problems, hence it works very well here.

Exercises 5.3

5.1. (a) $6\mathbf{i} + 6\mathbf{j}$; $-4\mathbf{i} + 4\mathbf{j}$,
(b) $2\mathbf{i}$; $-2\mathbf{j}$,
(c) $2\mathbf{i} - 2\mathbf{k}$; $-2\mathbf{j}$,
(d) $-\mathbf{i} + 4\mathbf{j} - \mathbf{k}$; $5\mathbf{i} - 2\mathbf{j} + 5\mathbf{k}$,
(e) $\mathbf{i} + \mathbf{j}$; $\mathbf{i} - \mathbf{j} + 2\mathbf{k}$.

5.2. Direction cosines are as follows:

(a) $\frac{1}{\sqrt{26}}, \frac{5}{\sqrt{26}}$; $\frac{5}{\sqrt{26}}, \frac{1}{\sqrt{26}}$

(b) $\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}$; $\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}$

(c) $\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}$; $\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}$

(d) $\frac{2}{3}, \frac{1}{3}, \frac{2}{3}$; $-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}$

(e) $\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}$; $\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}$.

5.3. $\alpha = -2$, $\beta = 1$, $\gamma = -3$

5.4. $A = 40.64°$, $B = 49.36°$, $C = 90°$.

5.5. $\theta = \dfrac{\pi}{2}, \dfrac{3\pi}{2}$.

5.6. $-1, -3, 1$.

5.7. (a) 67.38°, (b) 90°, (c) 70.5°, (d) 125.26°, (e) 90°.

5.8. (a) $p = -\dfrac{1}{2}$, (b) $p = 0$ or 4.

5.9. $\mathbf{j} + \mathbf{k}$.

5.10. From the definition of scalar product, $\hat{\mathbf{r}}_1 \cdot \hat{\mathbf{r}}_2 = |\hat{\mathbf{r}}_1||\hat{\mathbf{r}}_2| \cos(\alpha + \beta) = \cos(\alpha + \beta)$. However, using components, $\hat{\mathbf{r}}_1 \cdot \hat{\mathbf{r}}_2 = (\cos\beta, \sin\beta) \cdot (\cos\alpha, -\sin\alpha) = \cos\alpha \cos\beta + \sin\alpha\sin\beta$, whence $\cos(\alpha + \beta) = \cos\alpha\cos\beta + \sin\alpha\sin\beta$. Similarly, $\hat{\mathbf{r}}_1 \times \hat{\mathbf{r}}_2 = |\hat{\mathbf{r}}_1||\hat{\mathbf{r}}_2| \sin(\alpha + \beta)\mathbf{k}$

$$= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \cos\alpha & -\sin\alpha & 0 \\ \cos\beta & \sin\beta & 0 \end{vmatrix}, \text{ whence } \sin(\alpha + \beta)$$

$= \sin\alpha \cos\beta + \cos\alpha \sin\beta$.

5.11. $\mathbf{a} \times \mathbf{b} = 4\mathbf{i} + 3\mathbf{j} - 5\mathbf{k}$, $\mathbf{b} \times \mathbf{a} = -4\mathbf{i} - 3\mathbf{j} + 5\mathbf{k}$, $(\mathbf{a} - \mathbf{b}) \times (\mathbf{a} + \mathbf{b}) = 8\mathbf{i} + 6\mathbf{j} - 10\mathbf{k}$. $(\mathbf{a} - \mathbf{b}) \times (\mathbf{a} + \mathbf{b}) = \mathbf{a} \times \mathbf{a} - \mathbf{b} \times \mathbf{a} + \mathbf{a} \times \mathbf{b} - \mathbf{b} \times \mathbf{b} = 2\mathbf{a} \times \mathbf{b}$.

5.12. (a) 3,

(b) $\dfrac{\mathbf{F} \cdot (1, 2, 3)}{\sqrt{1 + 4 + 9}} = \dfrac{19}{\sqrt{14}}$,

(c) The work done is $\dfrac{\mathbf{F} \cdot (2, 5, 6)}{\sqrt{4 + 25 + 36}} = \dfrac{40}{\sqrt{65}}$.

5.13. Using the notation of Figure 5.15, A is $\mathbf{i} + \mathbf{k}$, B is $\mathbf{j} + \mathbf{k}$, C is $\mathbf{i} + \mathbf{j}$ and D is $\mathbf{i} + \mathbf{j} + \mathbf{k}$.

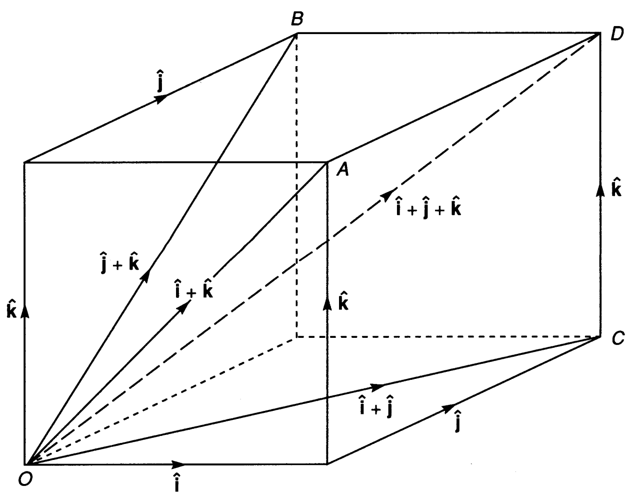


Figure 5.15 A cube and its diagonals.

That $\mathbf{i} + \mathbf{k} + \mathbf{j} + \mathbf{k} + \mathbf{i} + \mathbf{j} = 2(\mathbf{i} + \mathbf{j} + \mathbf{k})$ proves the result.

5.14. $\mathbf{a} = (1, -2, 1)$, $\mathbf{b} = (0, -2, 3)$ whence $\mathbf{b} - \mathbf{a} = (-1, 0, 2)$ and the line has equation $\mathbf{r} = (1 - \lambda, -2, 1 + 2\lambda)$. The plane that contains the points $(0, 0, 0)$, $(2, 4, 1)$ and $(4, 0, 2)$ has equation $\mathbf{r} \times (2, 4, 1) \cdot (4, 0, 2) = 0$, the condition that three vectors are coplanar. This expands and simplifies to $x = 2z$. The value of λ corresponding to a point in this plane is thus $1 - \lambda = 2 + 4\lambda$ or $\lambda = -\dfrac{1}{5}$.

The point of intersection is thus $\left(\dfrac{6}{5}, -2, \dfrac{3}{5}\right)$.

5.15. The result follows from the formula $\frac{1}{2}ab\sin C$ for the area of ΔABC which is

$\frac{1}{2}\left|\overrightarrow{PQ}\right|\left|\overrightarrow{PR}\right|\sin\theta = \frac{1}{2}\left|\overrightarrow{PQ} \times \overrightarrow{PR}\right|$ (see Figure 5.16).

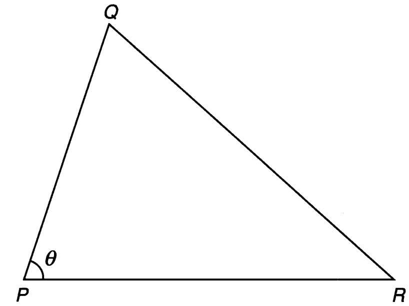


Figure 5.16 The triangle PQR.

(a) $\dfrac{1}{2}\sqrt{3}$, (b) $\dfrac{5}{2}\sqrt{5}$.

5.16. Given the vertices as position vectors, the sides are:

$\overrightarrow{AB} = (-3, 0, -3)$, $\overrightarrow{BC} = (4, 2, 4)$ and $\overrightarrow{CA} = (1, 2, -7)$.

Whence $\cos A = \dfrac{\overrightarrow{AB} \cdot \overrightarrow{AC}}{\left|\overrightarrow{AB}\right|\left|\overrightarrow{AC}\right|} = \dfrac{1}{3}\sqrt{3}$, and similarly,

$\cos C = \dfrac{1}{3}\sqrt{6}$. As $\overrightarrow{AB} \cdot \overrightarrow{BC} = -12 + 12 = 0$, B is a right angle.

5.17. $\dfrac{1}{\sqrt{50}}(5\mathbf{i} - 3\mathbf{j} + 4\mathbf{k})$.

5.18. The sides are $\mathbf{a} - \mathbf{b}$, $\mathbf{b} - \mathbf{c}$, $\mathbf{c} - \mathbf{a}$ so the area is

$\dfrac{1}{2}(\mathbf{a} - \mathbf{b}) \times (\mathbf{b} - \mathbf{c})$ which expands to

$\dfrac{1}{2}(\mathbf{a} \times \mathbf{b} + \mathbf{b} \times \mathbf{c} + \mathbf{c} \times \mathbf{a})$ as required.

5.19. Since $\hat{\mathbf{n}}$ bisects $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$ we have from the vector addition law, $\hat{\mathbf{a}} + \hat{\mathbf{b}} = k\hat{\mathbf{n}}$. Take the scalar product with $\hat{\mathbf{n}}$ to give $\hat{\mathbf{n}} \cdot \hat{\mathbf{a}} + \hat{\mathbf{n}} \cdot \hat{\mathbf{b}} = k$. Take the scalar product with $\hat{\mathbf{a}}$ then with $\hat{\mathbf{b}}$ to give the two equations $1 + \hat{\mathbf{a}} \cdot \hat{\mathbf{b}} = k\hat{\mathbf{a}} \cdot \hat{\mathbf{n}}$, and $1 + \hat{\mathbf{a}} \cdot \hat{\mathbf{b}} = k\,\hat{\mathbf{b}} \cdot \hat{\mathbf{n}}$, so $\hat{\mathbf{a}} \cdot \hat{\mathbf{n}} = \hat{\mathbf{b}} \cdot \hat{\mathbf{n}}$, and $k = 2(\hat{\mathbf{a}} \cdot \hat{\mathbf{n}})$. Substituting for k gives the result.

5.20. $\mathbf{a} \times (\mathbf{b} \times \mathbf{a}) = a^2\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{a}$, using the vector triple product. Whence $(\mathbf{a} \times (\mathbf{b} \times \mathbf{a})) \cdot \mathbf{b} = a^2b^2 - (\mathbf{a} \cdot \mathbf{b})^2$, 'dotting' with $\mathbf{b}$. The left-hand side can be interpreted as a scalar triple product so $\mathbf{b} \cdot \mathbf{a} \times (\mathbf{b} \times \mathbf{a}) = \mathbf{b} \times \mathbf{a} \cdot (\mathbf{b} \times \mathbf{a})$ and this is $(\mathbf{b} \times \mathbf{a})^2 = (\mathbf{a} \times \mathbf{b})^2$ which proves the result.

5.21. Given $\mathbf{x}^2 + \dfrac{\mathbf{b} \cdot \mathbf{x}}{a} + \dfrac{c}{a} = 0$ we complete the square to give

$\left(\mathbf{x} + \dfrac{\mathbf{b}}{2a}\right)^2 - \dfrac{\mathbf{b}^2}{4a^2} + \dfrac{c}{a} = 0$, and taking the square root

gives $\mathbf{x} = -\dfrac{\mathbf{b}}{2a} + \left(\dfrac{c}{a} - \dfrac{\mathbf{b}^2}{4a^2}\right)^{1/2}\hat{\mathbf{e}}$ where $\hat{\mathbf{e}}$ is an arbitrary unit vector.

5.22. The equation of the line in parametric form is: $\mathbf{r} = (2 + 6t)\mathbf{i} + (-3 + 2t)\mathbf{j} + (-1 + 3t)\mathbf{k}$ from which $t = \dfrac{x - 2}{6} = \dfrac{y + 3}{2} = \dfrac{z + 1}{3}$ as required. A corresponds to $t = 0$ and an arbitrary point P corresponds to $(2 + 6t, -3 + 2t, -1 + 3t)$. For AP to be 14 we require $36t^2 + 4t^2 + 9t^2 = (14)^2$, and so $t = \pm 2$ giving the two points $(14, 1, 5)$ and $(-10, -7, -7)$.

5.23. Figure 5.17 shows the set-up.

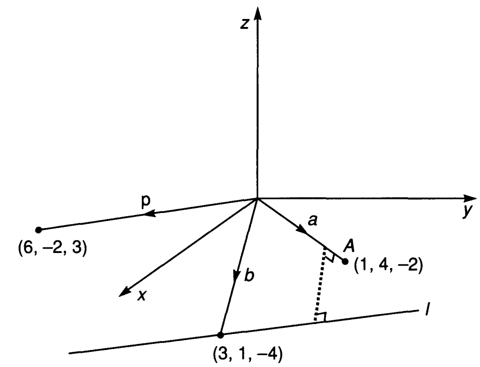


Figure 5.17 The line l is parallel to p; the dotted line is the required perpendicular distance.

$\mathbf{a} = (1, 4, -2)$, $\mathbf{b} = (3, 1, -4)$ and $\mathbf{p} = (6, -2, 3)$, so a line joining an arbitrary point on the line l to A is $\mathbf{r} - \mathbf{a}$ where $\mathbf{r} = \mathbf{b} + t\mathbf{p}$. This is to be perpendicular to $\mathbf{p}$ which implies $(\mathbf{r} - \mathbf{a}) \cdot \mathbf{p} = 0$ so $t = \dfrac{(\mathbf{a} - \mathbf{b}) \cdot \mathbf{p}}{\mathbf{p}^2} = \dfrac{12}{49}$ using the data.

Thus, the required distance is

$$\left|(2, -3, -2) - \frac{12}{49}(6, -2, 3)\right| = 3.71$$

5.24. In vector form the planes are $\mathbf{r} \cdot (2, 3, -4) = 1$ and $\mathbf{r} \cdot (3, 1, -2) = 2$. The vector perpendicular to the line of intersection is $(2, 3, -4) \times (3, 1, -2) = (-2, -8, -7)$. The plane is thus $\mathbf{r} \cdot (-2, -8, -7) = -c$ and to pass through $(2, -4, 5)$, $c = 7$, which gives the result.

5.25. The intersection of the two planes is given by $22y = 54 - 14z = 209 - 77x$ which can be considered to be the equation of the line of intersection. Writing the two planes in vector form $\mathbf{r} \cdot (3, -1, 1) = 12$, and $\mathbf{r} \cdot (1, 4, -2) = -5$ the perpendicular plane is $\mathbf{r} \cdot (3, -1, 1) \times (1, 4, -2) = k$ and if this is to pass through the point $(1, 2, -1)$, its equation is $-2x + 7y + 13z = -1$.

5.26. The cube of side a is shown in Figure 5.18 together with the diagonals FG and AE.

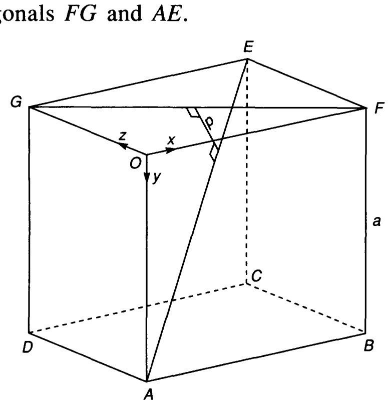


Figure 5.18 The cube of side a, the diagonals FG and AE, and the shortest distance between them, p.

Label the co-ordinates as follows:
F is $(a, 0, 0)$, A is $(0, a, 0)$, G is $(0, 0, a)$, B is $(a, a, 0)$, D is $(0, a, a)$, E is $(a, 0, a)$ and C is (a, a, a). From this we see that $\overrightarrow{FG} = \overrightarrow{FO} + \overrightarrow{OG} = -a\mathbf{i} + a\mathbf{k}$ and $\overrightarrow{EA} = a\mathbf{j} - a\mathbf{i} - a\mathbf{k}$. A line perpendicular to both is

$(\overrightarrow{FG}) \times (\overrightarrow{EA}) = -a^2(\mathbf{i} + 2\mathbf{j} + \mathbf{k})$, so a unit vector in this direction is $\dfrac{1}{\sqrt{6}}(\mathbf{i} + 2\mathbf{j} + \mathbf{k})$. $\overrightarrow{AF}$ is the vector $-a\mathbf{j} + a\mathbf{i}$ so

the required perpendicular distance is
$\dfrac{1}{\sqrt{6}}(\mathbf{i} + 2\mathbf{j} + \mathbf{k}).\overrightarrow{AF} = \dfrac{-a}{\sqrt{6}}$. As this must be positive, the distance is $a/\sqrt{6}$.

Figure 5.19 shows the triangle OEA.

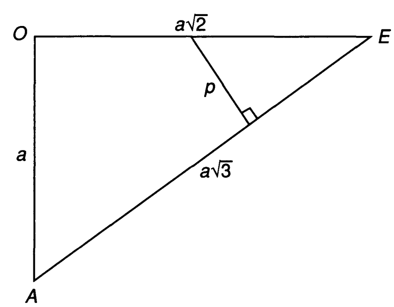


Figure 5.19 The triangle OEA and the shortest distance p.

The equations of the two diagonals are as follows:

$\overrightarrow{OA} + t\,\overrightarrow{AE} = (0, a, 0) + t(a, -a, a)$ and $\overrightarrow{AF} + s\overrightarrow{FG} = (0, -a, 0) + s(-a, 0, a)$.

5.27. The centre of the sphere is, say, (b, b, b). It touches the plane $x + y + z = a$ at the point $\left(\dfrac{a}{3}, \dfrac{a}{3}, \dfrac{a}{3}\right)$. However, the distance to the centre of the sphere from the origin is $b\sqrt{3}$ which means that the radius of the sphere must be $a\dfrac{\sqrt{3}}{3} - b\sqrt{3}$ from which $b = \dfrac{a\sqrt{3}}{(1 + \sqrt{3})}$ so the centre of the sphere is at the point $\dfrac{a}{6}(3 - \sqrt{3})(1, 1, 1)$.

5.28. By taking the scalar product with $\mathbf{v}$ derive the expression $\mathbf{v} \cdot \mathbf{r} = \gamma\mathbf{v} \cdot \mathbf{r}' + \gamma\mathbf{v}^2 t$; and using $1 + \gamma^2\dfrac{v^2}{c^2} = \gamma^2$ gives $t = \gamma t' + \gamma\dfrac{\mathbf{v} \cdot \mathbf{r}'}{c^2}$. Substituting for $\mathbf{v} \cdot \mathbf{r}$ and t in the expression $\mathbf{r}' = \mathbf{r} + \left[\dfrac{\gamma - 1}{\mathbf{v}^2}\mathbf{v} \cdot \mathbf{r} - \gamma t\right]\mathbf{v}$ gives after manipulation

$$\mathbf{r} = \mathbf{r}' + \left[\frac{\gamma - 1}{\mathbf{v}^2}\mathbf{v} \cdot \mathbf{r}' + \gamma t'\right]\mathbf{v}.$$

5.29. If $\mathbf{a}$ and $\mathbf{b}$ are the position vectors of the ends of a diameter, AB say, of a sphere, then $\mathbf{r} - \mathbf{a}$ and $\mathbf{r} - \mathbf{b}$ together with AB form a triangle APB where P has the position vector $\mathbf{r}$. The condition that P describes a sphere is the condition that the angle $A\hat{P}B$ is a right angle, that is $(\mathbf{r} - \mathbf{a}) \cdot (\mathbf{r} - \mathbf{b}) = 0$.

5.30. The lines are $\mathbf{r} = (9, 0, 5) + t(-4, 1, 1) = \mathbf{a} + t\mathbf{b}$ and $\mathbf{r} = (7, 8, 0) + s(2, 1, 1) = \mathbf{a}' + s\mathbf{b}'$. Using Examples 5.21 and 5.22 gives the perpendicular distance as

$$(\mathbf{a}' - \mathbf{a})\frac{\mathbf{b}' \times \mathbf{b}}{|\mathbf{b}' \times \mathbf{b}|} = (-2, 8, -5)\frac{1}{\sqrt{11}}(1, 1, -3) = \frac{14}{11}.$$

Exercises 6.3

6.1. Using the definition of derivative, we have

$$\frac{d}{dt}(\phi\mathbf{F}) = \lim_{\Delta t\to 0}\left\{\frac{\phi(t+\Delta t)\mathbf{F}(t+\Delta t) - \phi(t)\mathbf{F}(t)}{\Delta t}\right\}.$$ Adding and subtracting the term $\phi(t+\Delta t)\mathbf{F}(t)$ from the numerator then taking the limit gives the result.

6.2. $\mathbf{F}\cdot\frac{d\mathbf{F}}{dt} = F_1\frac{dF_1}{dt} + F_2\frac{dF_2}{dt} + F_3\frac{dF_3}{dt}$

$= \frac{1}{2}\frac{d}{dt}(F_1^2 + F_2^2 + F_3^2) = F\frac{dF}{dt}.$

6.3. $\frac{d\mathbf{L}}{dt} = \frac{d}{dt}\left(\mathbf{r}\times m\frac{d\mathbf{r}}{dt}\right) = \mathbf{r}\times\frac{d}{dt}\left(m\frac{d\mathbf{r}}{dt}\right) + \frac{d\mathbf{r}}{dt}\times m\frac{d\mathbf{r}}{dt}$

the second term of which is clearly zero. Hence

$\frac{d\mathbf{L}}{dt} = \mathbf{r}\times m\frac{d^2\mathbf{r}}{dt^2} = \mathbf{r}\times\mathbf{F} = \Gamma$. So $\mathbf{r}\times\mathbf{F} = \mathbf{0} \Rightarrow \mathbf{L} = \text{const}$.

6.4. $\mathbf{v} = (2t-6)\mathbf{i} + (3t^2 - 2t)\mathbf{j}\sqrt{2} + 2\mathbf{k}$,

$\mathbf{a} = 2\mathbf{i} + (6t-2)\mathbf{j}\sqrt{2}$,

$\mathbf{v}\cdot\mathbf{a} = (t-1)(18t^2+6) = 0$ when $t = 1$.

6.5 $\frac{d}{dt}\left[\frac{\mathbf{r}+\mathbf{a}}{\mathbf{r}^2+\mathbf{a}^2}\right] = \frac{\dot{\mathbf{r}}}{\mathbf{r}^2+\mathbf{a}^2} - \frac{2r\dot{r}(\mathbf{r}+\mathbf{a})}{(\mathbf{r}^2+\mathbf{a}^2)^2}$,

$\frac{d}{dt}\left[\frac{\mathbf{r}\times\mathbf{a}}{\mathbf{r}\cdot\mathbf{a}}\right] = \frac{(\mathbf{r}\cdot\mathbf{a})(\dot{\mathbf{r}}\times\mathbf{a}) - (\mathbf{r}\times\mathbf{a})(\dot{\mathbf{r}}\cdot\mathbf{a})}{(\mathbf{r}\cdot\mathbf{a})^2}.$

6.6. (a) $\mathbf{r} = \frac{1}{2}\mathbf{a}t^2 + \mathbf{A}t + \mathbf{B}$ where $\mathbf{A}$ and $\mathbf{B}$ are arbitrary constant vectors.

(b) Convert to $\frac{d^2\mathbf{r}}{dt^2} = \lambda\mathbf{a} + \frac{\mathbf{b}\times\mathbf{a}}{a^2}$ (see Example 5.17),

then integrate to give $\mathbf{r} = \frac{1}{2}\lambda\mathbf{a}t^2 + \frac{\mathbf{b}\times\mathbf{a}}{2a^2}t^2 + \mathbf{C}t + \mathbf{D}$ where $\mathbf{C}$ and $\mathbf{D}$ are arbitrary constant vectors.

6.7. (a) $\hat{\mathbf{T}} = \frac{(2t, 3t^2-1, 0)}{\sqrt{9t^4 - 2t^2 + 1}}$,

(b) $\hat{\mathbf{T}} = \frac{(a\cos t, -a\sin t, t)}{\sqrt{a^2+b^2}}$,

(c) $\hat{\mathbf{T}} = \frac{(1, 2t, 3t^2)}{\sqrt{9t^4 + 4t^2 + 1}}$.

6.8. Written in the form $\mathbf{r} = (1,1,1) + \lambda(1,2,1)$ and $\mathbf{r} = (0,0,2) + \mu(2,1,-4)$ the scalar product $(2,1,-4)\cdot(1,2,1) = 2+2-4 = 0$ shows that the lines are at right angles.

6.9. $\frac{d\mathbf{r}}{dt} = (a\cos t, -a\sin t, b) = (y, -x, b)$ as required.

6.10. For a plane curve, since $\hat{\mathbf{T}}$ and $\hat{\mathbf{N}}$ are in the plane of the curve, $\hat{\mathbf{B}}$ is a constant vector directed perpendicularly to the plane of the curve. The third Serret–Frenet formula $\frac{d\hat{\mathbf{B}}}{ds} = -\tau\hat{\mathbf{N}}$ thus implies $-\tau\hat{\mathbf{N}} = \mathbf{0}$, from which τ is zero.

6.11. Using the method of Example 6.8 we obtain:

$\hat{\mathbf{T}} = \left(-\frac{a}{c}\sin t, \frac{a}{c}\cos t, \frac{b}{c}\right)$, $\kappa = \frac{a}{c^2}$,

$\hat{\mathbf{N}} = (-\cos t, -\sin t, 0)$,

$\hat{\mathbf{B}} = \left(\frac{b}{c}\sin t, -\frac{b}{c}\cos t, \frac{b}{c}\right)$ and $\tau = \frac{b}{c^2}$

where $c^2 = a^2 + b^2$. In this example, $t = \frac{s}{c}$.

6.12. First of all $\frac{d\mathbf{r}}{d\theta} = (6\theta^2, 6\theta, 3)$

whence $\frac{ds}{d\theta} = \left|\frac{d\mathbf{r}}{d\theta}\right| = 3(4\theta^4 + 4\theta^2 + 1)^{1/2} = 3(2\theta^2+1)$

as required. The method of Example 6.8 then gives the following results:

$\hat{\mathbf{T}} = \frac{1}{(2\theta^2+1)}(2\theta^2, 2\theta, 1)$, $\kappa = \frac{2}{3}(2\theta^2+1)^{-2}$,

$\hat{\mathbf{N}} = \frac{1}{(2\theta^2+1)}(2\theta, 1-2\theta^2, -2\theta)$,

$\hat{\mathbf{B}} = \frac{1}{(2\theta^2+1)}(-1, 2\theta, -2\theta^2)$ and $\tau = -\frac{2}{3}(2\theta^2+1)^{-2}$.

From which $|\kappa| = |\tau| = \frac{2}{3}(2\theta^2+1)^{-2}$.

6.13. $\mathbf{r} = (2a\cos t, 2a\sin t, bt^2)$,

$\dot{\mathbf{r}} = (-2a\sin t, 2a\cos t, 2bt)$ and

$\ddot{\mathbf{r}} = (-2a\cos t, -2a\sin t, 2b)$

from which $\dot{\mathbf{r}}\cdot\ddot{\mathbf{r}} = 4b^2t$ follows at once. Taking cross products, we obtain

$\dot{\mathbf{r}}\times\ddot{\mathbf{r}} = (4ab\cos t + 4abt\sin t, -4abt\cos t + 4ab\sin t, 4a^2)$

and $|\dot{\mathbf{r}}\times\ddot{\mathbf{r}}| = \sqrt{16a^2b^2 + 16a^2b^2t^2 + 16a^4}$

$= 4a[a^2 + b^2 + b^2t^2]^{1/2}$.

Finally, $\hat{\mathbf{T}} = \frac{\dot{\mathbf{r}}}{|\dot{\mathbf{r}}|} = \frac{1}{\sqrt{a^2+b^2t^2}}(-a\sin t, a\cos t, bt)$.

6.14. $\mathbf{v} = \frac{d\mathbf{r}}{dt} = (U\sin\omega t, U\cos\omega t, V)$,

$\mathbf{v}\times\mathbf{B} = (BU\cos\omega t, -BU\sin\omega t, 0)$. Also,

$m\frac{d\mathbf{v}}{dt} = m(\omega U\cos\omega t, -\omega U\sin\omega t, 0)$. Thus

$m\frac{d\mathbf{v}}{dt} = q(\mathbf{v}\times\mathbf{B})$ provided $m\omega = qB$.

6.15. $\mathbf{r} = (x, y, z)$ thus $\frac{d\mathbf{r}}{ds} = (x', y', z') = \hat{\mathbf{T}}$ by definition. Hence

$\frac{d^2\mathbf{r}}{ds^2} = (x'', y'', z'') = \frac{d\hat{\mathbf{T}}}{ds}$, but the right-hand side is $\kappa\hat{\mathbf{N}}$ using the first of the Serret–Frenet formulae. Taking modulus thus gives $\kappa = \sqrt{(x'')^2 + (y'')^2 + (z'')^2}$ as required.

6.16. This problem demands extensive use of the Serret–Frenet formulae.

$\frac{d\mathbf{r}}{ds} = \hat{\mathbf{T}}$ and $\frac{d^2\mathbf{r}}{ds^2} = \frac{d\hat{\mathbf{T}}}{ds} = \kappa\hat{\mathbf{N}}$. Also,

$\frac{d^3\mathbf{r}}{ds^3} = \kappa\frac{d\hat{\mathbf{N}}}{ds} + \frac{d\kappa}{ds}\hat{\mathbf{N}} = \frac{d\kappa}{ds}\hat{\mathbf{N}} - \kappa\tau\hat{\mathbf{B}}$.

Thus, $\frac{d^2\mathbf{r}}{ds^2}\times\frac{d^3\mathbf{r}}{ds^3} = \kappa\hat{\mathbf{N}}\times\left(\frac{d\kappa}{ds}\hat{\mathbf{N}} - \kappa\tau\hat{\mathbf{B}}\right) = -\kappa^2\tau\hat{\mathbf{T}}$.

Taking the scalar product of this with $\frac{d\mathbf{r}}{ds} = \hat{\mathbf{T}}$ gives the result.

6.17. Using $d\mathbf{r} = \frac{\partial \mathbf{r}}{\partial u}du + \frac{\partial \mathbf{r}}{\partial v}dv$ gives

$$d\mathbf{r} \cdot d\mathbf{r} = ds^2 = \left(\frac{\partial \mathbf{r}}{\partial u}du + \frac{\partial \mathbf{r}}{\partial v}dv\right) \cdot \left(\frac{\partial \mathbf{r}}{\partial u}du + \frac{\partial \mathbf{r}}{\partial v}dv\right)$$

$$= \frac{\partial \mathbf{r}}{\partial u} \cdot \frac{\partial \mathbf{r}}{\partial u}du^2 + 2\frac{\partial \mathbf{r}}{\partial u} \cdot \frac{\partial \mathbf{r}}{\partial v}du dv + \frac{\partial \mathbf{r}}{\partial v} \cdot \frac{\partial \mathbf{r}}{\partial v}dv^2$$

which shows the first part.

If u and v are orthogonal, then $\frac{\partial \mathbf{r}}{\partial u} \cdot \frac{\partial \mathbf{r}}{\partial v} = 0$ and F is zero.

If F is zero, then $\frac{\partial \mathbf{r}}{\partial u} \cdot \frac{\partial \mathbf{r}}{\partial v} = 0$, which implies that u and v are orthogonal.

Exercises 7.3

7.1. $\nabla\phi = 2x(y^2 + z^2)\mathbf{i} + 2y(z^2 + x^2)\mathbf{j} + 2z(x^2 + y^2)\mathbf{k}$, so the unit normal at (1, 1, 1) is $\frac{1}{\sqrt{3}}(1, 1, 1)$ and that at (1, 0, −1) is $\frac{1}{\sqrt{2}}(1, 0, -1)$. It is just coincidence that the unit normals have the same form as the points themselves, this being a consequence of the form of $\nabla\phi$. The normal at $(a, 0, -a)$ is $\mathbf{n}_1 = 2a^3\mathbf{i} - 2a^3\mathbf{k}$ and the normal at $(a, 0, a)$ is $\mathbf{n}_2 = 2a^3\mathbf{i} + 2a^3\mathbf{k}$. It follows immediately that $\mathbf{n}_1 \cdot \mathbf{n}_2 = 0$ thus the normals are perpendicular.

7.2. (a) $\frac{\partial\phi}{\partial x} = nx^{n-1}$, $\frac{\partial\phi}{\partial y} = ny^{n-1}$ and $\frac{\partial\phi}{\partial z} = nz^{n-1}$, so the result follows.

(b) $\frac{\partial\phi}{\partial x} = ax^{a-1}y^bz^c$, $\frac{\partial\phi}{\partial y} = bx^ay^{b-1}z^c$ and $\frac{\partial\phi}{\partial z} = cx^ay^bz^{c-1}$ from which the result follows.

7.3. At (a, b, c), $\nabla\phi = \frac{2}{a}\mathbf{i} + \frac{2}{b}\mathbf{j} + \frac{2}{c}\mathbf{k}$ from which the direction cosines are $k(a, b, c)$ where $k = \left(\frac{1}{a^2} + \frac{1}{b^2} + \frac{1}{c^2}\right)^{1/2}$.

7.4. This result follows from the use of the quotient rule on the components

$$\nabla\left(\frac{\phi}{\psi}\right) = \frac{\partial}{\partial x}\left(\frac{\phi}{\psi}\right)\mathbf{i} + \frac{\partial}{\partial y}\left(\frac{\phi}{\psi}\right)\mathbf{j} + \frac{\partial}{\partial z}\left(\frac{\phi}{\psi}\right)\mathbf{k}.$$

7.5. (a) $\phi = \ln xyz$,

(b) $\phi = \sin xyz$,

(c) $\phi = x^2y^3z^4$,

(d) $\phi = xy + \frac{1}{2}x^2y^2z$.

7.6. Use $\phi = \frac{p_1x + p_2y + p_3z}{(x^2 + y^2 + z^2)^{3/2}}$ and differentiate to give:

$$\mathbf{E} = -\nabla\phi = -\frac{\mathbf{p}}{r^3} + \frac{3(\mathbf{p} \cdot \mathbf{r})\mathbf{p}}{r^5}$$

7.7. (a) $\frac{13}{\sqrt{29}}$,

(b) $-\frac{23}{\sqrt{45}}$,

(c) $\frac{116}{\sqrt{45}}$.

7.8. Maximum change is along the direction $\nabla T = aT_0(1 + cz + by)e^{ax}\mathbf{i} + bT_0e^{ax}\mathbf{j} + cT_0e^{ax}\mathbf{k}$ which at the origin is $T_0(a, b, c)$.

7.9. The unit normal is $\frac{1}{\sqrt{13}}(-2, 1, 3)$, found through $\nabla\phi$ in the usual way. The tangent plane is $\mathbf{r} \cdot (-2, 1, 3) = c$ where c is a constant. So the plane passing through the point $(1, -3, 2)$ is $-2x + y + 3z = 1$.

7.10. The angle between the surfaces is also the angle between the normals to the surfaces. We check that the point $(1, 2, -1)$ lies on both surfaces, then calculate both normals: $\nabla\phi = (6x, -2y, 2)$, $\nabla\psi = (y^2z - 3, 2xyz, xy^2 + 2z)$. At the point $(1, 2, -1)$ the normals are $(6, 4, 2)$ and $(1, -4, 6)$ whence the angle is

$$\cos^{-1}\left[\frac{2}{\sqrt{56}\sqrt{53}}\right] = 87.9°.$$

7.11. $\nabla \cdot \mathbf{F} = y + z$, $\nabla \times \mathbf{F} = -y\mathbf{i} - x\mathbf{k}$ and $\nabla(\nabla \cdot \mathbf{F}) = \mathbf{j} + \mathbf{k}$.

7.12. (a) $4xz - 4xyz + 6yz$,

(b) $2xy - xz + 2yz$,

(c) $a_1yze^{xyz} + a_2xze^{xyz} + a_3xye^{xyz}$.

7.13. (a) $(2z^4 + 2x^2y)\mathbf{i} + (3xz^2 + 4xyz)\mathbf{j} - 4xyz\mathbf{k}$,

(b) $\mathbf{0}$, since $\phi\nabla\phi = \nabla(\frac{1}{2}\phi^2)$, or use the formula $\nabla \times (\alpha\mathbf{A}) = \alpha\nabla \times \mathbf{A} + \nabla\alpha \times \mathbf{A}$ with $\alpha = \phi$, and $\mathbf{A} = \nabla\phi$.

(c) $\mathbf{0}$, using the determinant for expanding curl.

7.14. (a) Using the determinant for curl

$$\nabla \times (\phi\mathbf{A}) = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \phi A_1 & \phi A_2 & \phi A_3 \end{vmatrix} = \mathbf{i}\left(\frac{\partial}{\partial y}(\phi A_3) - \frac{\partial}{\partial z}(\phi A_2)\right) + \cdots$$

The identity follows by expanding the derivatives using the product rule then grouping the ϕ terms and the $\mathbf{A}$ terms.

(b) This is a particularly nasty formula to derive. The only practical way without recourse to computer algebra is to note that the $\mathbf{i}$ component of the left-hand side is $\frac{\partial}{\partial x}(A_1B_1 + A_2B_2 + A_3B_3)$, then to show that this is also the $\mathbf{i}$ component of the right-hand side. This is done term by term:

$$(\mathbf{B} \cdot \nabla)\mathbf{A} = \mathbf{i}\left(B_1\frac{\partial A_1}{\partial x} + B_2\frac{\partial A_1}{\partial y} + B_3\frac{\partial A_1}{\partial z}\right) + \cdots,$$

$$(\mathbf{A} \cdot \nabla)\mathbf{B} = \mathbf{i}\left(A_1\frac{\partial B_1}{\partial x} + A_2\frac{\partial B_1}{\partial y} + A_3\frac{\partial B_1}{\partial z}\right) + \cdots,$$

$$\mathbf{B} \times (\nabla \times \mathbf{A}) = \mathbf{i}\left(B_2\frac{\partial A_2}{\partial x} - B_2\frac{\partial A_1}{\partial y} - B_3\frac{\partial A_1}{\partial z} + B_3\frac{\partial A_3}{\partial x}\right) + \cdots,$$

$$\mathbf{A} \times (\nabla \times \mathbf{B}) = \mathbf{i}\left(A_2\frac{\partial B_2}{\partial x} - A_2\frac{\partial B_1}{\partial y} - A_3\frac{\partial B_1}{\partial z} + A_3\frac{\partial B_3}{\partial x}\right) + \cdots,$$

Adding these together reveals that all but the x derivatives cancel, and further these x derivatives combine to form $\frac{\partial}{\partial x}(A_1B_1 + A_2B_2 + A_3B_3)$, the left-hand side. We can argue

by symmetry that the **j** and **k** terms behave similarly, hence the identity is established.

(c) This is done as part (b). The left-hand side has **i** component

$$\nabla \times (\mathbf{A} \times \mathbf{B}) = \mathbf{i}\left(\frac{\partial}{\partial y}(A_1B_2 - B_1A_2) - \frac{\partial}{\partial z}(B_1A_3 - A_1B_3)\right) + \ldots$$

The right-hand side has the four terms

$$= \mathbf{i}\left(B_1\frac{\partial}{\partial x} + B_2\frac{\partial}{\partial y} + B_3\frac{\partial}{\partial z}\right)A_1 - \mathbf{i}\left(A_1\frac{\partial}{\partial x} + A_2\frac{\partial}{\partial y} + A_3\frac{\partial}{\partial z}\right)B_1$$

$$-\mathbf{i}B_1\left(\frac{\partial A_1}{\partial x} + \frac{\partial A_2}{\partial y} + \frac{\partial A_3}{\partial z}\right) + \mathbf{i}A_1\left(\frac{\partial B_1}{\partial x} + \frac{\partial B_2}{\partial y} + \frac{\partial B_3}{\partial z}\right) + \ldots$$

Once again, many terms cancel and the two **i** terms are shown to be the same. Symmetry then establishes the identity. Put $\mathbf{A} = \mathbf{B} = \mathbf{u}$ in part (b) to give $\nabla(\mathbf{u}\cdot\mathbf{u}) = 2(\mathbf{u}\cdot\nabla)\mathbf{u} - 2\mathbf{u} \times \nabla \times \mathbf{u}$ from which $(\mathbf{u}\cdot\nabla)\mathbf{u} = \nabla(\frac{1}{2}\mathbf{u}^2) - \mathbf{u} \times \nabla \times \mathbf{u}$. For irrotational flow, the second term on the right is zero and so the potential is $\frac{1}{2}\mathbf{u}^2$.

7.15. Since $\nabla\cdot(\nabla\phi \times \nabla\psi) = \nabla\psi\cdot\nabla\times\nabla\phi - \nabla\phi\cdot\nabla\times\nabla\psi = 0$ the combination $\nabla\phi \times \nabla\psi$ is solenoidal and the result is established.

7.16. From the definition of divergence and curl we can immediately write that

$$\nabla\phi\cdot\nabla\psi \times \nabla\Omega = \begin{vmatrix} \dfrac{\partial\phi}{\partial x} & \dfrac{\partial\phi}{\partial y} & \dfrac{\partial\phi}{\partial z} \\ \dfrac{\partial\psi}{\partial x} & \dfrac{\partial\psi}{\partial y} & \dfrac{\partial\psi}{\partial z} \\ \dfrac{\partial\Omega}{\partial x} & \dfrac{\partial\Omega}{\partial y} & \dfrac{\partial\Omega}{\partial z} \end{vmatrix} = \frac{\partial(\phi, \psi, \Omega)}{\partial(x, y, z)}$$

Hence if ϕ, ψ and Ω are functionally related, this determinant must be zero. This is equivalent to the three gradients $\nabla\phi$, $\nabla\psi$ and $\nabla\Omega$ lying in the same plane.

7.17. By direct differentiation

$$\nabla^2(\phi\psi) = \frac{\partial^2}{\partial x^2}(\phi\psi) + \frac{\partial^2}{\partial y^2}(\phi\psi) + \frac{\partial^2}{\partial z^2}(\phi\psi)$$
$$= \phi\,\frac{\partial^2\psi}{\partial x^2} + 2\,\frac{\partial\phi}{\partial x}\,\frac{\partial\psi}{\partial x} + \frac{\partial^2\phi}{\partial x^2} + \ldots$$

so that

$$\nabla^2(\phi\psi) = \phi\nabla^2\psi + 2\nabla\phi\cdot\nabla\psi + \psi\nabla^2\phi.$$

7.18. Substitution and use of the rules of differentiation establish that the given forms of **E** and **H** satisfy Maxwell's equations as stated.

7.19. (See 7.18).

7.20. Demanding that the flow satisfies $\nabla\cdot\mathbf{u} = 0$ implies $a + l = 0$. Demanding that the flow is also irrotational, that is $\nabla \times \mathbf{u} = \mathbf{0}$ gives $k - b = 0$. Hence

$$\mathbf{u} = (ax + by)\mathbf{i} + (bx - ay)\mathbf{j} = \frac{\partial\phi}{\partial x}\,\mathbf{i} + \frac{\partial\phi}{\partial y}\mathbf{j}.$$

Integration gives the result.

7.21. Differentiation gives:

$$\frac{\partial\mathbf{r}}{\partial\rho} = -\cos\alpha\,\cos\beta\mathbf{i} - \cos\alpha\,\sin\beta\mathbf{j} + \sin\alpha\mathbf{k}$$

$$\frac{\partial\mathbf{r}}{\partial\alpha} = \rho\sin\alpha\,\cos\beta\mathbf{i} + \rho\sin\alpha\sin\beta\mathbf{j} + \rho\,\cos\alpha\mathbf{k}$$

$$\frac{\partial\mathbf{r}}{\partial\beta} = -(a - \rho\cos\alpha)\sin\beta\mathbf{i} + (a - \rho\cos\alpha)\cos\beta\mathbf{j}$$

Hence $h_1 = \left|\dfrac{\partial\mathbf{r}}{\partial\rho}\right| = 1$, $h_2 = \left|\dfrac{\partial\mathbf{r}}{\partial\alpha}\right| = \rho$ and

$h_3 = \left|\dfrac{\partial\mathbf{r}}{\partial\beta}\right| = (a - \rho\cos\alpha)$. Also,

$$\hat{\mathbf{e}}_1 = -\cos\alpha\cos\beta\mathbf{i} - \cos\alpha\sin\beta\mathbf{j} + \sin\alpha\mathbf{k}$$
$$\hat{\mathbf{e}}_2 = \sin\alpha\cos\beta\mathbf{i} + \sin\alpha\sin\beta\mathbf{j} + \cos\alpha\mathbf{k}$$
$$\hat{\mathbf{e}}_3 = -\sin\beta\mathbf{i} + \cos\beta\mathbf{j}$$

are easily tested to be orthogonal.

7.22. In orthogonal curvilinear co-ordinates curl takes the form:

$$\nabla \times \mathbf{A} = \begin{vmatrix} h_1\hat{\mathbf{e}}_1 & h_2\hat{\mathbf{e}}_2 & h_3\hat{\mathbf{e}}_3 \\ \dfrac{\partial}{\partial u_1} & \dfrac{\partial}{\partial u_2} & \dfrac{\partial}{\partial u_3} \\ h_1A_1 & h_2A_2 & h_3A_3 \end{vmatrix} \text{ so}$$

$$\nabla\cdot(\nabla \times \mathbf{A}) = \begin{vmatrix} \dfrac{\partial}{\partial u_1} & \dfrac{\partial}{\partial u_2} & \dfrac{\partial}{\partial u_3} \\ \dfrac{\partial}{\partial u_1} & \dfrac{\partial}{\partial u_2} & \dfrac{\partial}{\partial u_3} \\ h_1A_1 & h_2A_2 & h_3A_3 \end{vmatrix}$$

which is zero for the same reason as it is in Cartesian co-ordinates.

$$\nabla \times \nabla\phi = \begin{vmatrix} h_1\hat{\mathbf{e}}_1 & h_2\hat{\mathbf{e}}_2 & h_3\hat{\mathbf{e}}_3 \\ \dfrac{\partial}{\partial u_1} & \dfrac{\partial}{\partial u_2} & \dfrac{\partial}{\partial u_3} \\ \dfrac{\partial\phi}{\partial u_1} & \dfrac{\partial\phi}{\partial u_2} & \dfrac{\partial\phi}{\partial u_3} \end{vmatrix} = \mathbf{0} \text{ by directly multiplying out.}$$

7.23. $\nabla\cdot(\phi\mathbf{e}_\rho) = \dfrac{1}{\rho(a - \rho\cos\alpha)}\,\dfrac{\partial}{\partial\rho}(\phi\rho(a - \rho\cos\alpha))$. So if this is zero, we have $\phi\rho(a - \rho\cos\alpha) = f(\alpha)$ where the right-hand side is an arbitrary function of the co-ordinate α.

Hence $\phi = \dfrac{f(\alpha)}{\rho(a - \rho\cos\alpha)}$. On $\rho = a$, $\phi = (\sin^2\frac{1}{2}\alpha)^{-1}$,

hence $\dfrac{f(\alpha)}{a(a - a\cos\alpha)} = \dfrac{1}{\sin^2\frac{1}{2}\alpha}$ or $f(\alpha) = \dfrac{a\cdot 2a\sin^2\frac{1}{2}\alpha}{\sin^2\frac{1}{2}\alpha} = 2a^2$.

Whence the solution is $\phi = \dfrac{2a^2}{\rho(a - \rho\cos\alpha)}$.

Exercises 8.3

8.1. (a) $\int_0^\pi a^4\cos\theta\sin^4\theta d\theta = 0$, (b) $9\int_0^1 t^8 dt = 1$.

(c) $\int_0^{2\pi}(a^2\sin\theta + b^2\cos\theta)d\theta = \pi(a^2 + b^2)$.

8.2. (a) 2, (b) $2\pi^2$, (c) $2\pi^2$. (b) and (c) are equal because the end points are the same and the integral is independent of the path since $\mathbf{r} = \nabla(\frac{1}{2}\mathbf{r}^2)$.

8.3. (a) 2, (b) 2. $\phi = x^2y + z^2$.

8.4. (a) $\frac{8}{15}\mathbf{i} + \frac{6}{35}\mathbf{j} + \frac{16}{567}\mathbf{k}$, (b) $\frac{20 + 3\sqrt{2}}{15}$,

(c) $\frac{9 - 2\sqrt{2}}{18}\mathbf{i} - \frac{8}{9}\mathbf{j} + \frac{2\sqrt{2} - 1}{2}\mathbf{k}$.

8.5. $\int_C \mathbf{A}.\mathrm{d}\mathbf{r} = \int_C A_1\mathrm{d}x + A_2\mathrm{d}y + A_3\mathrm{d}z = [A_1x + A_2y + A_3z]_C$ since $\mathbf{A}$ is a constant vector. This is zero since x, y and z have the same values at both 'ends' of C.

8.6. $\int_C \mathbf{r}\mathrm{d}s = (0, 0, 2\pi b(a^2 + b^2)^{1/2})$ and

$\int_C \mathbf{r} \times \mathrm{d}\mathbf{r} = (0, 2\pi ab, 2\pi a^2)$.

8.7. -6π.

8.8. (a) $\frac{8}{3}$, (b) 4.

8.9. $-\frac{a^4}{2}\pi\mathbf{k}$.

8.10. (a) $2\pi a$, (b) 4, (c) $2 + \sqrt{2}$. In fact $\int_C \hat{\mathbf{T}} \cdot \mathrm{d}\mathbf{r} = \int_C \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s} \cdot \mathrm{d}\mathbf{r} = \int_C \mathrm{d}s$ which is the perimeter of C.

8.11. The integral is $\mathbf{B} = \int_0^{2\pi} \frac{-z(b\cos\theta\,\mathbf{i} + a\sin\theta\,\mathbf{j}) - ab\mathbf{k}}{(a^2\cos^2\theta + b^2\sin^2\theta + z^2)^{1/2}}\,\mathrm{d}\theta$.

Exercises 9.3

9.1. (a) $\frac{2}{3}$, (b) $\frac{1}{6}$, (c) $2\sin 1 - \sin 2$; the fact that $0 \le \iint_D \sin(x + y)\,\mathrm{d}x\mathrm{d}y \le 1$ follows since the modulus of the integrand is always less than one.

(d) $\frac{4}{5}$.

9.2. (a)

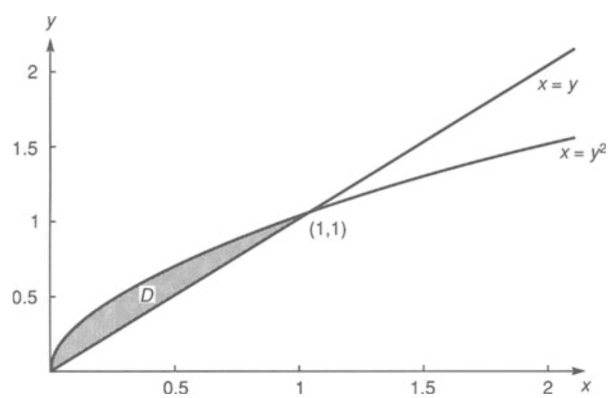


Figure 9.21

$\frac{2}{27}$.

(b)

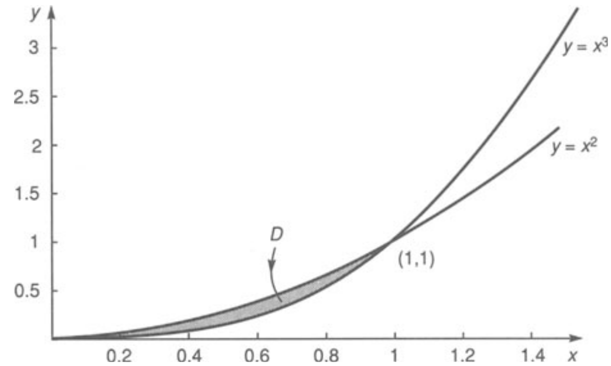


Figure 9.22

$\frac{9}{280}$.

(c)

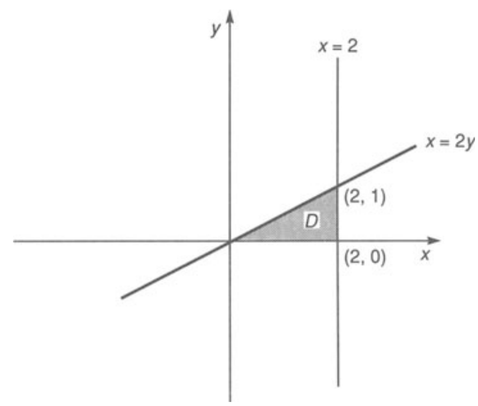


Figure 9.23

$\frac{1}{4}(e^4 - 1)$.

(d)

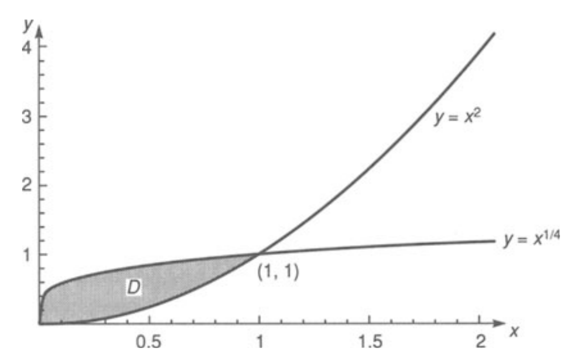


Figure 9.24

$\frac{9}{70}$.

9.3. (a)

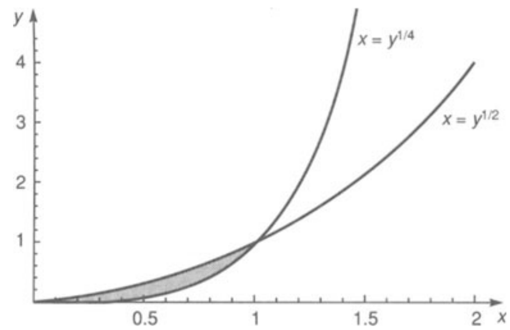


Figure 9.25

$$\int_0^1 \int_{y^{1/2}}^{y^{1/4}} f(x, y)\mathrm{d}x\mathrm{d}y.$$

(b)

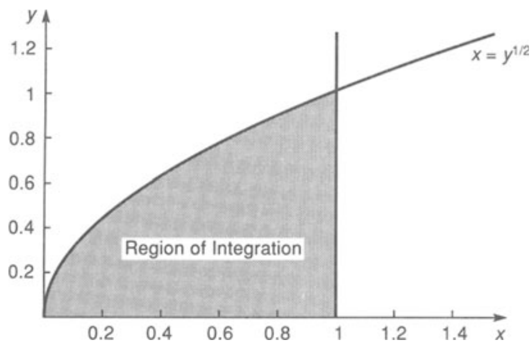


Figure 9.26

$$\int_0^1 \int_{y^{1/2}}^{1} f(x, y)\mathrm{d}x\mathrm{d}y.$$

(c)

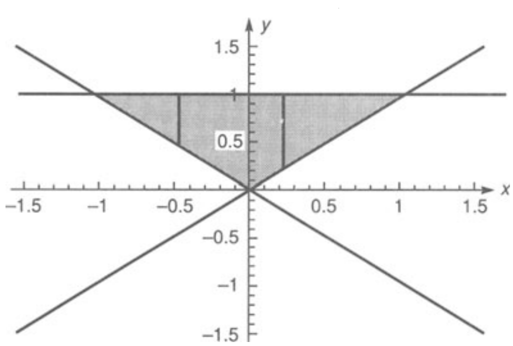


Figure 9.27 The triangular region of integration and two typical vertical strips.

$$\int_{-1}^{0} \int_{-x}^{1} f(x, y)\mathrm{d}y\mathrm{d}x + \int_0^1 \int_x^1 f(x, y)\mathrm{d}y\mathrm{d}x.$$

(d)

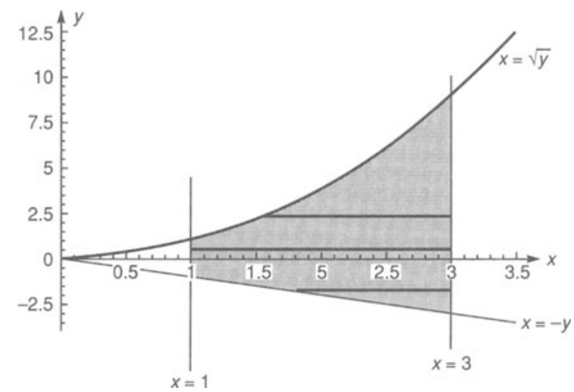


Figure 9.28 The shaded region is the domain of the integral, and three typical horizontal strips are shown.

$$\int_{-3}^{-1} \int_{-y}^{3} f(x, y)\mathrm{d}x\mathrm{d}y + \int_{-1}^{1} \int_{1}^{3} f(x, y)\mathrm{d}x\mathrm{d}y + \int_{1}^{9} \int_{\sqrt{y}}^{3} f(x, y)\mathrm{d}x\mathrm{d}y.$$

9.4. (a)

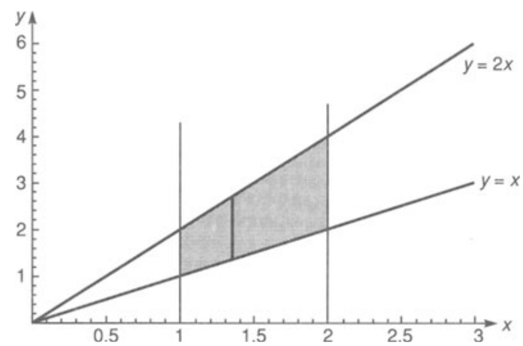


Figure 9.29 The region of integration (shaded). Integration with respect to y first.

$$\ln 2\left(\tan^{-1}2 - \frac{\pi}{4}\right).$$

(b)

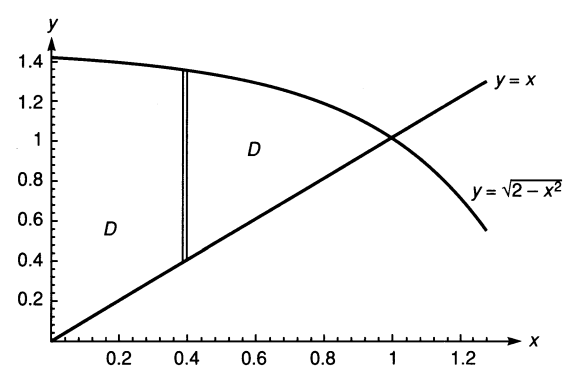


Figure 9.30 The region of integration D, and a typical vertical strip.

$$\frac{\sqrt{2}}{2}.$$

9.5. (a) $\displaystyle\int_0^1\int_{\sqrt{x}}^1 \sin\left(\frac{y^3+1}{2}\right) \mathrm{d}y\mathrm{d}x = \int_0^1\int_0^{y^2} \sin\left(\frac{y^3+1}{2}\right) \mathrm{d}x\mathrm{d}y =$

$\frac{2}{3}(\cos\frac{1}{2} - \cos 1)$,

(b) $\displaystyle\int_0^1\int_{x^2}^1 \frac{x^3}{\sqrt{x^4+y^2}}\mathrm{d}y\mathrm{d}x = \int_0^1\int_0^{\sqrt{y}} \frac{x^3}{\sqrt{x^4+y^2}}\,\mathrm{d}x\mathrm{d}y = \frac{1}{8}(\sqrt{2}-1)$,

(c) $\displaystyle\int_0^1\int_0^{\cos^{-1}y} e^{\sin x}\,\mathrm{d}x\mathrm{d}y = \int_0^{\pi/2}\int_0^{\cos x} e^{\sin x}\,\mathrm{d}y\mathrm{d}x = e - 1$,

(d) $\displaystyle\int_0^1\int_x^1 x^2e^{y^4}\,\mathrm{d}y\mathrm{d}x = \int_0^1\int_0^y x^2e^{y^4}\,\mathrm{d}x\mathrm{d}y = \frac{1}{12}(e-1)$.

9.6. The required area is 4 × shaded area

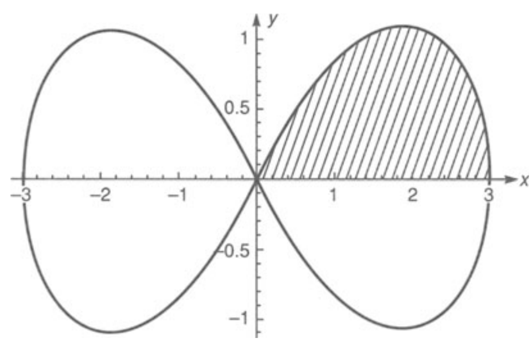


Figure 9.31 The lemniscate of Bernoulli with the region of integration shaded.

$$= 4\int_0^{\pi/4}\int_0^{3\cos 2\theta} R\mathrm{d}R\mathrm{d}\theta = \frac{9}{4}\pi.$$

9.7 $\frac{\pi}{12}$.

9.8. The integral becomes $\displaystyle\int_0^{\pi/2}\frac{2\cos 2\theta}{(1+\cos\theta)^2}\mathrm{d}\theta = \frac{1}{3}(6\pi - 20)$,

either by computer algebra or by using the substitution $t = \tan\frac{1}{2}\theta$.

9.9. The integral becomes $\displaystyle\int_0^{\pi/4}\int_0^1 2\ln R\mathrm{d}R\mathrm{d}\theta = -\frac{\pi}{2}$.

9.10. (a) For this transformation the Jacobian

$\dfrac{\partial(r,\theta)}{\partial(x,y)} = \dfrac{1}{4(\cos\theta\sin\theta)^{1/2}}$, hence the integral is

$$\int_0^{\pi/2}\int_0^1 \frac{(r\cos\theta)^{3/2}(1-r^2)}{4(\cos\theta\sin\theta)^{1/2}}\,\mathrm{d}r\mathrm{d}\theta = \frac{2}{45}\int_0^{\pi/2}\frac{\cos\theta}{\sqrt{\sin\theta}}\mathrm{d}\theta = \frac{4}{45}.$$

(b) The integral is $\displaystyle\int_0^{2\pi}\int_a^b 2R\ln R\mathrm{d}R\mathrm{d}\theta = \pi(b^2 - a^2)$.

(c) For this transformation the Jacobian $\dfrac{\partial(u,v)}{\partial(x,y)} = \dfrac{x+y}{x^2}$,

hence the integral becomes $\displaystyle\int_0^1\int_y^{2-y}\frac{x+y}{x^2}e^{x+y}\,\mathrm{d}x\mathrm{d}y =$

$\displaystyle\int_0^1\int_0^2 e^u\mathrm{d}u\mathrm{d}v = e^2 - 1$, and the triangle in the x–y plane transforms to the rectangle in the u–v plane as shown in Figure 9.32.

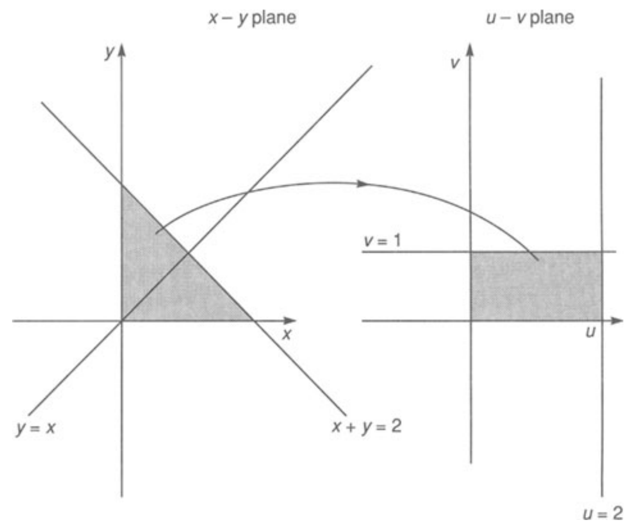


Figure 9.32 The shaded areas map to each other.

9.11. Using Green's Theorem the integral becomes

$$\int_0^a\int_0^a 4y\mathrm{d}y\mathrm{d}x = 2a^3$$

9.12. Using Green's Theorem

$$\oint_C e^x\sin y\mathrm{d}x + e^x\cos y\mathrm{d}x = \iint_S(e^x\cos y - e^x\cos y)\mathrm{d}x\mathrm{d}y = 0.$$

In terms of a potential, if $e^x\sin y = \dfrac{\partial\phi}{\partial x}$ and $e^x\cos y = \dfrac{\partial\phi}{\partial y}$ then $\phi = e^x\sin y$ and $\oint_C \nabla\phi\cdot\mathrm{d}\mathbf{s} = 0$.

9.13. Integrate this directly since

$$\int_V e^{-r^2}\mathrm{d}V = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{\infty}^{\infty} e^{-(x^2+y^2+z^2)}\mathrm{d}x\mathrm{d}y\mathrm{d}z = \left(\int_{-\infty}^{\infty}e^{-t^2}\mathrm{d}t\right)^3 = \pi^{3/2}.$$

9.14. (a) $\frac{8}{3}$, (b) 104, (c) $\frac{53}{105}$.

9.15.

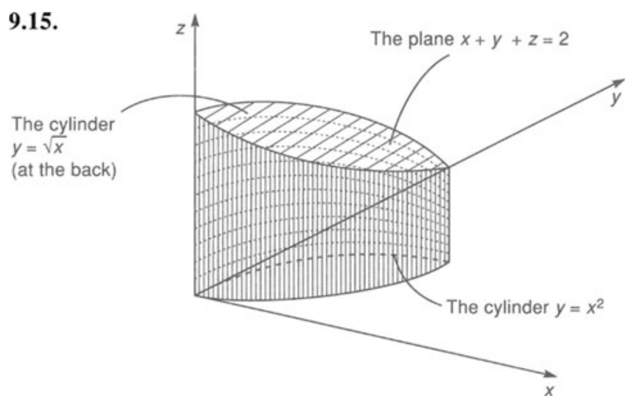


Figure 9.33

$$\int_0^1\int_{x^2}^{\sqrt{x}}\int_0^{2-x-y}\mathrm{d}z\mathrm{d}y\mathrm{d}x = \frac{11}{20}.$$

9.16. $6\pi - 2$.

9.17. $\frac{2\pi a^3}{3}(1 - \cos\alpha)$, and when $\alpha = \frac{\pi}{2}$, the volume is $\frac{4}{3}\pi a^3$.

9.18. $\frac{16}{5}\pi\sqrt{2}$.

9.19. The formulae follow immediately from writing $\bar{\mathbf{r}} = \dfrac{\int_D \mathbf{r}dA}{\int_D dA}$ in plane polar co-ordinates.

(a) For the cardioid, $\bar{x} = \frac{5}{6}\pi a$, $\bar{y} = 0$ (by symmetry).

(b) For the quarter lemniscate, computer algebra is useful for evaluating the two integrals $\int_0^{\pi/4}\cos^{3/2}2\theta\cos\theta d\theta$ and $\int_0^{\pi/4}\cos^{3/2}2\theta\sin\theta d\theta$; the results give

$$\bar{x} = \frac{1}{4}\pi a,\ \bar{y} = \left(\frac{1}{2}\ln(1 + \sqrt{2}) + \frac{7\sqrt{2}}{3}\right)a.$$

9.20. The volume is given by the triple integral

$$\int_{-\sqrt{3}}^{\sqrt{3}}\int_0^{2\pi}\int_1^{\sqrt{4-z^2}} RdRd\theta dz = 4\pi\sqrt{3}.$$

9.21. $\alpha a^2 r(\cos\theta\sin\theta)^{\alpha-1}$.

9.22. The integral transforms to $\frac{1}{2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{e^{-u^2}}{1+v^2}dudv = \frac{1}{2}\pi^{3/2}$.

9.23. The moment of inertia is given by the triple integral

$$\int_0^{2\pi}\int_0^h\int_0^{za/h} R^3dzdRd\theta = \frac{3}{10}ma^2$$ where m is the mass of the cone.

9.24. The volume of the ice cream cone is given by the integrals

$$\int_0^{2\pi}\int_0^{\sqrt{3}/2}\int_0^{z/\sqrt{3}} RdRdzd\theta + \int_0^{2\pi}\int_{\sqrt{3}/2}^{1}\int_0^{\sqrt{1-z^2}} RdRdzd\theta =$$

$$\frac{1}{3}\pi(2 - \sqrt{3}).$$

Similarly, the moment of inertia is the sum

$$\int_0^{2\pi}\int_0^{\sqrt{3}/2}\int_0^{z/\sqrt{3}} \rho R^3dRdzd\theta + \int_0^{2\pi}\int_{\sqrt{3}/2}^{1}\int_0^{\sqrt{1-z^2}} \rho R^3dRdzd\theta$$ where ρ is the density which is messy but straightforward to evaluate. In terms of the mass of the ice cream cone m, the answer is

$$\frac{286 - 147\sqrt{3}}{320(2 - \sqrt{3})}m.$$

Exercises 10.3

10.1. (a) 108, (b) $\frac{2\pi a^2}{3}$, (c) $\frac{81}{2}$.

10.2. (a) $\frac{8}{3}\pi a^3$, (b) $4\pi a^3$, (c) $3a^3$, (d) 320π.

(All of these answers can be verified using Gauss's Flux Theorem, see Chapter 11.)

10.3. Since $z - f(x, y) = 0$ is the equation of the surface S, we deduce that the unit normal to S takes the form

$$\hat{\mathbf{n}} = \left(-\frac{\partial f}{\partial x}\mathbf{i} - \frac{\partial f}{\partial y}\mathbf{j} + \mathbf{k}\right)\Bigg/\sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 + 1}.$$

Using the fact that area $= \int_S dS = \int_S \hat{\mathbf{n}}\cdot d\mathbf{S} = \int_R \dfrac{dR}{|\hat{\mathbf{n}}\cdot\hat{\mathbf{k}}|}$

then gives the result.

10.4. Use of projection gives

$$\int_S \mathbf{r}\cdot d\mathbf{S} = \int_R (2x^2 + 2y^2 + z)dR = \int_0^{2\pi}\int_0^1 (R^2 + 1)RdRd\theta$$

$$= \frac{3\pi}{2},$$ where we have $z = 1 - x^2 - y^2$ on R.

10.5. Using direct evaluation means parameterising the sphere which leads to

$$\int_S (x + y + z)dS =$$

$$\int_0^{\pi/2}\int_0^{\pi/2} (\sin\theta\cos\lambda - \sin\theta\sin\lambda + \cos\lambda)\sin\theta d\theta d\lambda$$

which is evaluated straightforwardly to be $\frac{3\pi}{4}$.

10.6. Evaluating the same integral as is in Exercise 10.5 but by projection leads to

$$\int_S (x + y + z)dS = \int_R (x + y + z)\frac{dR}{\cos\theta} =$$

$$\int_0^{\pi/2}\int_0^1\left(\frac{R(\cos\theta + \sin\theta)}{\sqrt{1 - R^2}} + 1\right)RdRd\theta.$$

This is evaluated reasonably straightforwardly to be $\frac{3\pi}{4}$. This method is longer than that of the last example.

10.7. $4\pi\lambda$, hence independent of the radius of the sphere.

10.8. $\frac{4\pi\mu}{a}$ hence no longer independent of the radius of the sphere.

10.9. We use the result that $dS = h_u h_v du dv$ where $h_u = \left|\frac{\partial \mathbf{r}}{\partial u}\right|$ and $h_v = \left|\frac{\partial \mathbf{r}}{\partial v}\right|$. (See Chapter 7.) $\frac{\partial \mathbf{r}}{\partial u} = \cos v\mathbf{i} + \sin v\mathbf{j}$, and $\frac{\partial \mathbf{r}}{\partial v} = -u\sin v\mathbf{i} + u\cos v\mathbf{j} + b\mathbf{k}$ whence

$$\int_S \sqrt{x^2 + y^2}dS = \int_0^1\int_0^{2\pi} u\sqrt{u^2 + b^2}dvdu = \frac{2\pi}{3}\sqrt{(1 + b^2)^{3/2} - b^3}.$$

10.10. When the rain comes straight down, the calculation is as follows:

$$\mathbf{F}\cdot\hat{\mathbf{n}} = \frac{z}{r} \text{ and } \hat{\mathbf{n}}\cdot\hat{\mathbf{k}} = \frac{z}{r} \text{ thus } \int_S \mathbf{F}\cdot d\mathbf{S} = \int_R dxdy = \pi.$$

If the rain is at a slant, the calculation proceeds in this fashion:

$\mathbf{F} = -\frac{\hat{\mathbf{i}}}{\sqrt{2}} - \frac{\hat{\mathbf{k}}}{\sqrt{2}}$, therefore $\mathbf{F}\cdot\hat{\mathbf{n}} = -\frac{1}{\sqrt{2}}\left(\frac{x}{r} - \frac{z}{r}\right)$, whence

$$\int_S \mathbf{F}\cdot d\mathbf{S} = -\int_R \frac{1}{\sqrt{2}}\left(\frac{x}{r} - \frac{z}{r}\right)\frac{r}{z}dR =$$

$$-\frac{1}{\sqrt{2}}\int_0^{2\pi}\int_0^1 \frac{R\cos\theta + R\sin\theta}{R}RdRd\theta$$ which has a value $\pi/\sqrt{2}$. This gives some insight into the inaccuracy of home-made rain gauges.

10.11. The area of a curved survace is $A(S) = \int_S dS$. Using results from curvilinear co-ordinates (Chapter 7), we have $dS = \left|\frac{\partial \mathbf{r}}{\partial u}\times\frac{\partial \mathbf{r}}{\partial v}\right|dudv$, whence since $\frac{\partial \mathbf{r}}{\partial u} = \frac{\partial x}{\partial u}\mathbf{i} + \frac{\partial y}{\partial u}\mathbf{j} + \frac{\partial z}{\partial u}\mathbf{k}$ and $\frac{\partial \mathbf{r}}{\partial v} = \frac{\partial x}{\partial v}\mathbf{i} + \frac{\partial y}{\partial v}\mathbf{j} + \frac{\partial z}{\partial v}\mathbf{k}$ we have

$$\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} & \frac{\partial z}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} & \frac{\partial z}{\partial v} \end{vmatrix}$$

$$= \frac{\partial(y, z)}{\partial(u, v)}\mathbf{i} + \frac{\partial(z, x)}{\partial(u, v)}\mathbf{j} + \frac{\partial(x, y)}{\partial(u, v)}\mathbf{k} \text{ and so}$$

$\left|\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}\right| = \sqrt{\left(\frac{\partial(y, z)}{\partial(u, v)}\right)^2 + \left(\frac{\partial(z, x)}{\partial(u, v)}\right)^2 + \left(\frac{\partial(x, y)}{\partial(u, v)}\right)^2}$ from which the result follows.

(a) For the cone, $\frac{\partial \mathbf{r}}{\partial u} = \cos v\mathbf{i} + \sin v\mathbf{j} + \mathbf{k}$ and $\frac{\partial \mathbf{r}}{\partial v} = -u\sin v\mathbf{i} + u\cos v\mathbf{j}$ and so (by evaluating the cross product and taking the modulus, it's quicker than using Jacobians) we have that the area of the cone is thus $\int_S dS = \int_0^1\int_0^{2\pi} u\sqrt{2}\,dv\,du = \pi\sqrt{2}$.

(b) For the helicoid, $\frac{\partial \mathbf{r}}{\partial u} = \cos v\mathbf{i} + \sin v\mathbf{j}$ and $\frac{\partial \mathbf{r}}{\partial v} = -u\sin v\mathbf{i} + u\cos v\mathbf{j} + \mathbf{k}$ so proceeding as before, the area of the helicoid is given by

$\int_S dS = \int_0^1\int_0^{2\pi}\sqrt{v^2 + 1}\,du\,dv = 2\pi\int_0^1\sqrt{v^2 + 1}\,dv = \pi(\sqrt{2} + \ln(1 + \sqrt{2}))$.

10.12. Writing $x = u$, $y = v$, $z = f(u, v)$ as a valid parameterisation of the surface S, we can derive the normal $\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}$ as $-\frac{\partial f}{\partial u}\mathbf{i} - \frac{\partial f}{\partial v}\mathbf{j} + \mathbf{k}$. The result then follows since $\int_S \mathbf{F}\cdot d\mathbf{S} = \int_R \mathbf{F}\cdot\left(\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}\right)dx\,dy$ upon writing $u = x$ and $v = y$.

Exercises 11.3

11.1. (a) 108π, (b) $3a^3$.

11.2. For $\mathbf{F} = \mathbf{r}$ the flux consists of the two parts:

$$\int_{\text{discs}} z\,dS + \int_{\substack{\text{curved}\\\text{part}}} (x^2 + y^2)\,dS = 2 \times \pi + 2 \times 2\pi = 6\pi.$$

Using the Divergence Theorem, $\nabla\cdot\mathbf{r} = 3$ so

$$\int_S \mathbf{F}\cdot d\mathbf{S} = \int_V \nabla\cdot\mathbf{F}\,dV = 3 \times \text{Volume} = 6\pi.$$

The second method is easier.

11.3. We have already shown that $\int_S (x^2\mathbf{i} + y^2\mathbf{j} + z^2\mathbf{k})\cdot d\mathbf{S} = \pi$ where S is the surface of the sphere $x^2 + y^2 + z^2 = 1$. We thus need to find $\int_V \nabla\cdot(x^2\mathbf{i} + y^2\mathbf{j} + z^2\mathbf{k})\,dV$ where V is the sphere $x^2 + y^2 + z^2 \le 1$. Using spherical polar coordinates this volume integral becomes

$$\int_0^1\int_0^\pi\int_0^{2\pi} 2r(\cos\lambda\sin\theta + \sin\theta\sin\lambda + \cos\theta)\,r^2\sin\theta\,dr\,d\theta\,d\lambda$$
$$= \tfrac{1}{2}\int_0^\pi\int_0^{2\pi}(\cos\lambda\sin\theta\sin\theta + \sin^2\theta\sin\lambda + \sin\theta\cos\theta)\,d\theta\,d\lambda$$
$$= \tfrac{1}{2}\int_0^\pi \pi(\sin\lambda + \cos\lambda)\,d\lambda = \pi.$$ Gauss's Theorem is thus confirmed.

11.4. Using the Divergence Theorem, $\nabla\cdot\mathbf{F} = 7x^2z^2$, so the flux is

$$\int_S \mathbf{F}\cdot d\mathbf{S} = \int_V 7x^2z^2\,dV = 7\int_0^\pi\int_0^{2\pi}\int_0^1 r^6\cos^2\theta\sin\theta\sin^2\phi\cos^2\phi\,dr\,d\theta\,d\phi$$

which integrates to $\frac{\pi}{6}$.

11.5. If $\mathbf{F} = \nabla\phi$ then $\nabla\cdot\mathbf{F} = \nabla^2\phi = 0$ since ϕ is harmonic. If $\nabla\cdot\mathbf{F} = 0$, then $\int_V \nabla\cdot\mathbf{F}\,dV = 0 = \int_S \mathbf{F}\cdot d\mathbf{S}$ for all closed surfaces S. Thus $\mathbf{F}\cdot d\mathbf{S}$ is an exact differential, $d\phi$ say. We thus have $\mathbf{F}\cdot d\mathbf{S} = F_1dx + F_2dy + F_3dz = d\phi$. Using the chain rule thus gives $F_1 = \frac{\partial\phi}{\partial x}$, $F_2 = \frac{\partial\phi}{\partial y}$, $F_3 = \frac{\partial\phi}{\partial z}$ so $\mathbf{F} = \nabla\phi$.

11.6. This needs no further comment.

11.7. Let $\mathbf{F} = \phi\mathbf{a}$ where ϕ is a scalar and $\mathbf{a}$ is an arbitrary but constant vector. Using the vector identity $\nabla\cdot\mathbf{F} = \nabla\cdot(\phi\mathbf{a}) = \mathbf{a}\cdot\nabla\phi$, Gauss's Divergence Theorem becomes $\int_V \nabla\cdot\mathbf{F}\,dV = \int_V \mathbf{a}\cdot\nabla\phi\,dV = \mathbf{a}\cdot\int_V \nabla\phi\,dV$ which is equal to $\int_S \mathbf{F}\cdot d\mathbf{S} = \int_S \mathbf{a}\phi\cdot d\mathbf{S} = \mathbf{a}\cdot\int_S \phi\,d\mathbf{S}$. We thus have $\mathbf{a}\cdot\left[\int_V \nabla\phi\,dV - \int_S \phi\,d\mathbf{S}\right] = 0$ and since $\mathbf{a}$ is arbitrary, the result follows.

11.8. Suppose that the function ϕ has an extremum at the point C, and take the origin of co-ordinates to be at C. Surround C by a small sphere of radius ε which lies entirely within the volume V. The fact that ϕ is harmonic together with Gauss's Theorem applied to the sphere then implies $\int_0^{2\pi}\int_0^\pi \frac{\partial\phi}{\partial r}\sin\theta\,d\theta\,d\lambda = 0$. Multiplying this result by dr and integrating from $r = 0$ to $r = \varepsilon$ yields $\int_0^\varepsilon\int_0^{2\pi}\int_0^\pi \frac{\partial\phi}{\partial r}\sin\theta\,d\theta\,d\lambda\,dr = 0$ or $\int_0^{2\pi}\int_0^\pi(\phi(\varepsilon) - \phi(0))\sin\theta\,d\theta\,d\lambda = 0$. Now, $\phi(0)$ is the extreme value that does not depend on either θ or λ, hence $\int_0^{2\pi}\int_0^\pi \phi(\varepsilon)\sin\theta\,d\theta\,d\lambda = 4\pi\phi(0)$. Multiplying this by ε^2 then integrating between 0 and ε gives $\phi(0) = \frac{1}{V_\varepsilon}\int_{V_\varepsilon}\phi\,dV$ where V_ε denotes the small sphere. This expression implies that $\phi(0)$ is the average value over the sphere. It cannot therefore be an extreme value and the result is proved.

11.9. Let $\mathbf{A} = \mathbf{a} \times \mathbf{F}$ where $\mathbf{a}$ is a constant vector, we have the identity $\nabla(\mathbf{a} \times \mathbf{F}) = \mathbf{F}\cdot\nabla \times \mathbf{a} - \mathbf{a}\cdot\nabla \times \mathbf{F} = -\mathbf{a}\cdot\nabla \times \mathbf{F}$. Apply Gauss's Divergence Theorem to obtain $-\mathbf{a}\cdot\int_V \nabla \times \mathbf{F}\,dV = \int_S \mathbf{a} \times \mathbf{F}\cdot d\mathbf{S} = \mathbf{a}\cdot\int_S \mathbf{F} \times d\mathbf{S}$ using a property of the scalar triple product. The result follows since $\mathbf{a}$ is an arbitrary vector.

11.10. (a) With $\mathbf{F} = (2x - y)\mathbf{i} - yz^2\mathbf{j} - y^2z\mathbf{k}$, $\nabla \times \mathbf{F} = \mathbf{k}$, whence $\int_S \nabla \times \mathbf{F}\cdot d\mathbf{S} = \int_S \mathbf{k}\cdot d\mathbf{S} = \int_R dx\,dy = \pi a^2$, where R is the disc $x^2 + y^2 \le a^2$, being the projection of the hemisphere on to the x–y plane. Directly, we have $\int_C \mathbf{F}\cdot d\mathbf{r} = \int_0^{2\pi}(-2a^2\cos\theta\sin\theta + a^2\sin^2\theta)\,d\theta = \pi a^2$.

(b) $\mathbf{F} = xy\mathbf{i} + yz\mathbf{j} + zx\mathbf{k}$, $\nabla \times \mathbf{F} = 6\mathbf{k}$ so that

$$\int_S \nabla \times \mathbf{F}\cdot d\mathbf{S} = 6\int_S \mathbf{k}\cdot d\mathbf{S} = 6\int_R dx\,dy = 6 \times \frac{\pi}{4} = \frac{3\pi}{2}.$$

Directly, $\mathbf{F} = -\frac{3}{2}\sin\theta\,\mathbf{i} + \frac{3}{2}\cos\theta\,\mathbf{j} + \frac{1}{4}\mathbf{k}$ and $\mathbf{dr} = -\frac{1}{2}\sin\theta\,\mathrm{d}\theta\mathbf{i} + \frac{1}{2}\cos\theta\,\mathrm{d}\theta\mathbf{j}$, whence $\int_C \mathbf{F}\cdot\mathbf{dr} = 2\pi\left(-\frac{3}{4}\right) = -\frac{3\pi}{2}$.

11.11. For this problem, an attempt to draw the curve C has been made in Figure 11.6.

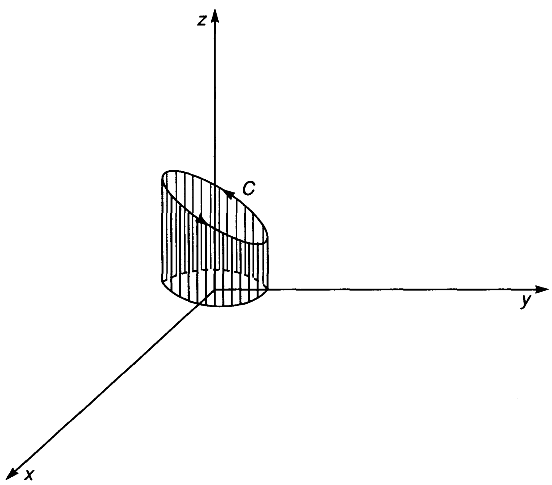


Figure 11.6 The intersection of the plane $z + y = a^2$ with the cylinder $x^2 + y^2 = b^2$ in the curve C.

$\nabla \times \mathbf{F} = -y\mathbf{i} - z\mathbf{j} - x\mathbf{k}$, but the area over which the surface integral is taken is shown shaded in Figure 11.6, and this is a 'sawn off' cylinder. The use of projection is not possible since the cylinder is wrapped around the z-axis. Instead we use parameterisation. The unit normal to the curved part of S is $\hat{\mathbf{n}} = \frac{1}{b}(x\mathbf{i} + y\mathbf{j})$, and noting that the contribution to $\int_S \nabla \times \mathbf{F}\cdot\mathbf{dS}$ from the flat disc on the $z = 0$ plane is zero, we have that $\int_S \nabla \times \mathbf{F}\cdot\mathbf{dS} = -\frac{1}{b}\int_S y(x+z)\mathrm{d}S$. On S, $y = b\sin t$, $x = b\cos t$, $z = z$ so that $\mathrm{d}S = b\mathrm{d}z\mathrm{d}t$ and so $-\frac{1}{b}\int_S y(x+z)\mathrm{d}S = -\int_0^{2\pi}\int_0^{a^2-b\sin t}(b^2\sin t\cos t + zb\sin t)\mathrm{d}z\mathrm{d}t = \pi a^2b^2$. This can be checked by direct evaluation of $\int_C \mathbf{F}\cdot\mathbf{dr}$. On C
$y = b\sin t$, $x = b\cos t$, $z = a^2 - b\sin t$ and the integrand $\mathbf{F}\cdot\mathbf{dr}$ is given by

$$\mathbf{F}\cdot\mathbf{dr} = -b^3\sin^2 t\cos t\mathrm{d}t + b^2\sin t\cos t(a^2 - b\sin t)\mathrm{d}t - b^2\cos^2 t(a^2 - b\sin t)\mathrm{d}t$$

and this integrated from 0 to 2π is clearly $-a^2b^2\int_0^{2\pi}\cos^2 t\mathrm{d}t = -\pi a^2b^2$. The negative sign results from the direction that C is traversed. In this case, direct evaluation is perhaps easier!

11.12. $\mathbf{P} = \mathbf{E}\times\mathbf{H}$ thus using the identity $\nabla\cdot(\mathbf{E}\times\mathbf{H}) = \mathbf{H}\cdot\nabla\times\mathbf{E} - \mathbf{E}\cdot\nabla\times\mathbf{H}$ in Gauss's Divergence Theorem gives $\int_S \mathbf{E}\times\mathbf{H}\cdot\mathbf{dS} = \int_V \nabla\cdot(\mathbf{E}\times\mathbf{H})\mathrm{d}V$, and expanding the integrand then using the two Maxwell's equations $\nabla\times\mathbf{E} = 0$ and $\nabla\times\mathbf{H} = \mathbf{J}$ gives the result.

Appendix A: Conjugate Harmonic Functions

In this appendix an outline is given of the topic of conjugate function theory. This subject belongs to Chapter 7, and some of the problems in that chapter are put into context by it. Consider a vector field $\mathbf{F}$ that satisfies the field equations:

$$\nabla \times \mathbf{F} = \mathbf{0}$$
$$\nabla \cdot \mathbf{F} = 0.$$

From the first of these, $\mathbf{F}$ is conservative (or irrotational) so the existence of a potential ϕ such that

$$\mathbf{F} = \nabla\phi$$

is assured. From the equation $\nabla \cdot \mathbf{F} = 0$, ϕ is harmonic, that is

$$\nabla^2 \phi = 0$$

If, in addition, it is assumed that ϕ is a function only of the two variables x and y, and not dependent on z, with $\mathbf{F} = F_1\mathbf{i} + F_2\mathbf{j}$, we have

$$\nabla \cdot \mathbf{F} = 0 \Rightarrow \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} = 0$$

This is satisfied by the function ψ where

$$F_1 = \frac{\partial \psi}{\partial y},\ F_2 = -\frac{\partial \psi}{\partial x}$$

provided ψ has continuous second-order partial derivatives.

The equation $\nabla \times \mathbf{F} = \mathbf{0}$ then implies

$$\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ F_1 & F_2 & 0 \end{vmatrix} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ \dfrac{\partial \psi}{\partial y} & -\dfrac{\partial \psi}{\partial x} & 0 \end{vmatrix} = 0$$

or

$$\mathbf{k}\left(-\frac{\partial^2 \psi}{\partial x^2} - \frac{\partial^2 \psi}{\partial y^2} \right) = \mathbf{0}$$

Hence we have the following equations valid for ϕ and ψ:

$$\nabla^2 \phi = 0$$
$$\nabla^2 \psi = 0$$

together with

$$\frac{\partial\phi}{\partial x} = \frac{\partial\psi}{\partial y}, \frac{\partial\phi}{\partial y} = -\frac{\partial\psi}{\partial x}$$

These last two equations arise from the two forms of the components of the vector field **F**:

$$\mathbf{F} = F_1\mathbf{i} + F_2\mathbf{j} = \frac{\partial\phi}{\partial x}\mathbf{i} + \frac{\partial\phi}{\partial y}\mathbf{j} = \frac{\partial\psi}{\partial y}\mathbf{i} - \frac{\partial\psi}{\partial x}\mathbf{j}$$

and are often referred to as the Cauchy–Riemann equations. The functions ϕ and ψ that are both harmonic and obey the Cauchy–Riemann equations are called *conjugate harmonic functions*. Conjugate harmonic functions play a key role in inviscid fluid dynamics where ϕ is the fluid potential and ψ is the stream function. They are also encountered in electromagnetism where the lines $\phi(x, y) =$ const. are equipotential lines and the lines $\psi(x, y) =$ const. are the electric or magnetic force lines.

The sets of lines $\phi(x, y) =$ const. and $\psi(x, y) =$ const. intersect at right angles, as can be seen from the following argument:

The normal to the line $\phi(x, y) =$ const. is of course $\nabla\phi$, and the normal to the line $\psi(x, y) =$ const. is of course $\nabla\psi$. Now,

$$\begin{aligned}\nabla\phi.\nabla\psi &= \frac{\partial\phi}{\partial x}\frac{\partial\psi}{\partial x} + \frac{\partial\phi}{\partial y}\frac{\partial\psi}{\partial y} \\ &= \frac{\partial\phi}{\partial x}\frac{\partial\phi}{\partial y} - \frac{\partial\phi}{\partial y}\frac{\partial\phi}{\partial x} \\ &= 0\end{aligned}$$

Therefore the normals are at right angles and so the tangents to the curves must be orthogonal.

For those familiar with complex variables, functions of a complex variable $z = x + iy$, $i = \sqrt{-1}$ give a rich source of conjugate harmonic functions. This is because the real and imaginary parts of a regular complex function are conjugate harmonic functions:

$$f(z) = f(x + iy) = \phi(x, y) + i\psi(x, y)$$

The proof of this is outside the scope of this text, but the following table gives some conjugate harmonic functions generated in this way:

$f(z)$	$\phi(x, y)$	$\psi(x, y)$
z^2	$x^2 - y^2$	$2xy$
$\dfrac{1}{z}$	$\dfrac{x}{x^2 + y^2}$	$-\dfrac{y}{x^2 + y^2}$
e^z	$e^x\cos y$	$e^x\sin y$
$\sin z$	$\sin x \cosh y$	$\cos x \sinh y$

Finally, some mention is made of conjugate harmonic functions in plane polar co-ordinates. In plane polar co-ordinates (r, θ), the Cauchy–Riemann equations are:

$$\frac{\partial\phi}{\partial r} = \frac{1}{r}\frac{\partial\psi}{\partial\theta}, \frac{1}{r}\frac{\partial\phi}{\partial\theta} = -\frac{\partial\psi}{\partial r}$$

Examples of conjugate harmonic functions in plane polar co-ordinates are: $r^n\cos n\theta$, $r^n\sin n\theta$ (n a constant); $\ln r$, θ.

Appendix B: Vector Calculus

B.1 Vector Identities

In this appendix a list of common vector identities is given. It is assumed that ϕ and ψ are scalar-valued functions with continuous second-order partial derivatives and that $\mathbf{F}$ and $\mathbf{G}$ are vector-valued functions with continuous second-order partial derivatives:

$$\nabla(\phi + \psi) = \nabla\phi + \nabla\psi$$
$$\nabla \cdot (\mathbf{F} + \mathbf{G}) = \nabla \cdot \mathbf{F} + \nabla \cdot \mathbf{G}$$
$$\nabla \times (\mathbf{F} + \mathbf{G}) = \nabla \times \mathbf{F} + \nabla \times \mathbf{G}$$
$$\nabla \cdot (\phi\mathbf{F}) = (\nabla\phi) \cdot \mathbf{F} + \phi\nabla \cdot \mathbf{F}$$
$$\nabla \times (\phi\mathbf{F}) = (\nabla\phi) \times \mathbf{F} + \phi\nabla \times \mathbf{F}$$
$$\nabla \cdot (\mathbf{F} \times \mathbf{G}) = \mathbf{G} \cdot \nabla \times \mathbf{F} - \mathbf{F} \cdot \nabla \times \mathbf{G}$$
$$\nabla \times (\mathbf{F} \times \mathbf{G}) = (\mathbf{G} \cdot \nabla)\mathbf{F} - \mathbf{G}(\nabla \cdot \mathbf{F}) - (\mathbf{F} \cdot \nabla)\mathbf{G} + \mathbf{F}(\nabla \cdot \mathbf{G})$$
$$\nabla(\mathbf{F} \cdot \mathbf{G}) = (\mathbf{G} \cdot \nabla)\mathbf{F} + (\mathbf{F} \cdot \nabla)\mathbf{G} + \mathbf{G} \times (\nabla \times \mathbf{F}) + \mathbf{F} \times (\nabla \times \mathbf{G})$$
$$\nabla \times (\nabla\phi) = \mathbf{0}$$
$$\nabla \cdot (\nabla \times \mathbf{F}) = 0$$
$$\nabla \times (\nabla \times \mathbf{F}) = \nabla(\nabla \cdot \mathbf{F}) - \nabla^2\mathbf{F}$$

where in any co-ordinate system other than Cartesian, the last equation is to be taken as the definition of ∇^2. ∇^2 on the other hand has been defined in terms of orthogonal curvilinear co-ordinates in Chapter 7 (see the next section).

B.2 Curvilinear Co-ordinates

If $\hat{\mathbf{e}}_1$, $\hat{\mathbf{e}}_2$ and $\hat{\mathbf{e}}_3$ are orthogonal unit vectors in the directions defined by mutually orthogonal co-ordinates u, v and w then:

$$\mathrm{d}\mathbf{r} = \frac{\partial \mathbf{r}}{\partial u}\mathrm{d}u + \frac{\partial \mathbf{r}}{\partial v}\mathrm{d}v + \frac{\partial \mathbf{r}}{\partial w}\mathrm{d}w = h_1\mathrm{d}u\hat{\mathbf{e}}_1 + h_2\mathrm{d}v\hat{\mathbf{e}}_2 + h_3\mathrm{d}w\hat{\mathbf{e}}_3$$

so that the quantities: h_1, h_2, h_3 and $\hat{\mathbf{e}}_1$, $\hat{\mathbf{e}}_2$, $\hat{\mathbf{e}}_3$ define particular curvilinear systems. The quantities $\nabla\phi$, $\nabla \cdot \mathbf{F}$, $\nabla \times \mathbf{F}$ and $\nabla^2\phi$ are given in terms of curvilinear co-ordinates as follows:

$$\nabla\phi = \frac{1}{h_1}\frac{\partial\phi}{\partial u}\hat{\mathbf{e}}_1 + \frac{1}{h_2}\frac{\partial\phi}{\partial v}\hat{\mathbf{e}}_2 + \frac{1}{h_3}\frac{\partial\phi}{\partial w}\hat{\mathbf{e}}_3$$

$$\nabla \cdot \mathbf{F} = \frac{1}{h_1h_2h_3}\left\{\frac{\partial}{\partial u}(h_2h_3F_1) + \frac{\partial}{\partial v}(h_3h_1F_2) + \frac{\partial}{\partial w}(h_1h_2F_3)\right\}$$

$$\nabla \times \mathbf{F} = \frac{1}{h_1h_2h_3}\begin{vmatrix} h_1\hat{\mathbf{e}}_1 & h_2\hat{\mathbf{e}}_2 & h_3\hat{\mathbf{e}}_3 \\ \dfrac{\partial}{\partial u} & \dfrac{\partial}{\partial v} & \dfrac{\partial}{\partial w} \\ h_1F_1 & h_2F_2 & h_3F_3 \end{vmatrix}$$

$$\nabla^2\phi = \frac{1}{h_1h_2h_3}\left\{\frac{\partial}{\partial u}\left[\frac{h_2h_3}{h_1}\frac{\partial\phi}{\partial u}\right] + \frac{\partial}{\partial v}\left[\frac{h_3h_1}{h_2}\frac{\partial\phi}{\partial v}\right] + \frac{\partial}{\partial w}\left[\frac{h_1h_2}{h_3}\frac{\partial\phi}{\partial w}\right]\right\}$$

The following table gives some h_1, h_2, h_3 for well known and perhaps less well known co-ordinate systems:

Title	h_1	h_2	h_3
Cylindrical Polars (R, θ, z)	1	R	1
Spherical Polars (r, θ, λ)	1	r	$r \sin \theta$
Parabolic Cylinder (u, v, z)	$\sqrt{u^2 + v^2}$	$\sqrt{u^2 + v^2}$	1
Paraboloidal (u, v, ϕ)	$\sqrt{u^2 + v^2}$	$\sqrt{u^2 + v^2}$	uv
Elliptic Cylinder (u, v, z)	$a\sqrt{\sin^2 u + \sinh^2 v}$	$a\sqrt{\sin^2 u + \sinh^2 v}$	1
Bipolar (u, v, z)	$\dfrac{a}{(\cosh v - \cos u)}$	$\dfrac{a}{(\cosh v - \cos u)}$	1

In the above table, for the elliptic cylinder co-ordinates, $(-a, 0)$ and $(a, 0)$ are the foci of the ellipses (u constant) and hyperbolae (v constant) that constitute the co-ordinate system. For the bipolar co-ordinates, the points $(-a, 0)$ and $(a, 0)$ are the centres of the two systems of circles that form the co-ordinates v = constant; the u co-ordinates are then circles that pass through these points.

Bibliography

In the course of producing this text, several published works have been consulted but none have been specifically cited. Below is a by no means exhaustive list of books that the reader might find relevant to the study of *Advanced Calculus*. The comments that follow each are my own personal opinion.

Durrant, A. V. (1996) *Vectors in Physics and Engineering*, 288pp. Chapman & Hall, London.
Good text for those who like the engineering approach. The examples favour units and numbers.

Dyke, P. P. G. (1995) *Mechanics (Work Out)*, 215pp. Macmillan, Basingstoke, UK.
The companion to the present text with a little overlap in Chapter 6. Hard not to recommend it!

Etgen, G. J. (1995) *Calculus (one and several variables) Salas and Hille*, 1370pp. John Wiley, New York.
This is one of many very thick calculus books that are brilliantly produced with wonderful diagrams, many in colour and in 3-D. This one is my personal favourite and is a seventh revision of a classic text by Salas and Hille. It is very good for the preliminary calculus of Chapter 1 (and before), but do not pay full price for it!

Gilbert, J. (1991) *Guide to Mathematical Methods*, 309pp. Macmillan, Basingstoke, UK.
This is here because it contains my favourite treatment of max. and min. of functions of two variables. Also covers the material of Chapters 2, 5 and 9.

Hirst, A. E. (1995) *Vectors in 2 or 3 Dimensions*, 134pp. Arnold, London.
A title in the 'Modular Mathematics' series, therefore brief. The treatment is more pure than the present text, but it is readable.

Lewis, P. E. and Ward, J. P. (1989) *Vector Analysis for Engineers and Scientists*, 406pp. Addison-Wesley, Wokingham, UK.
A more mathematical approach than Durrant and proceeding to a higher level. Highly recommended; slightly more applied than Marsden and Tromba.

Marsden, J. E. and Tromba, A. J. (1976) *Vector Calculus*, 655pp. W. H. Freeman, New York.
This is an excellent textbook that covers almost all of the topics (but not optimisation). There are later reprintings and perhaps a new edition. Highly recommended, but be prepared for some rigour.

Spiegel, M. R. (1959) *Vector Analysis*, 225pp. Schaum, McGraw-Hill, New York.
Perhaps the direct competitor to Spiegel (1963), this has dated but is still an excellent source book for both solved and unsolved problems.

Spiegel, M. R. (1963) *Advanced Calculus*, 384pp. Schaum, McGraw-Hill, New York.
This is the standard 'problem solver'. Now a little dated, and covering a wider class of problems than the present text (such as Series, Fourier Series, Complex Variable). Worth looking at.

Index